Дж. Н. Шайв

Физические свойства и конструкции полупроводниковых приборов

ГОСЭНЕРГОИЗДАТ

Дж. Н. ШАЙВ

ФИЗИЧЕСКИЕ СВОЙСТВА И КОНСТРУКЦИИ ПОЛУПРОВОДНИКОВЫХ ПРИБОРОВ

Перевод с английского
под редакцией
д. т. н., проф. *Г. А. ТЯГУНОВА*

ГОСУДАРСТВЕННОЕ ЭНЕРГЕТИЧЕСКОЕ ИЗДАТЕЛЬСТВО
МОСКВА 1963 ЛЕНИНГРАД

ЭС-5-4

В книге описаны физические свойства и элементарная теория полупроводниковых приборов; рассматриваются основные конструктивные принципы, а также некоторые эксплуатационные и схемные свойства этих приборов.

Книга предназначена для инженерно-технических работников, работающих над созданием и применением полупроводниковых приборов, и студентов высших учебных заведений

THE PROPERTIES, PHYSICS
AND DESIGN OF SEMICONDUCTOR DEVICES

BY J. N. SHIVE

D van Nostrand Co, Inc.
1959

6Ф2.13
Ш 17 Шайв Дж. Физические свойства и конструкции полупроводниковых приборов (перевод с английского под редакцией д. т. н., проф. Г. А. Тягунова), М. — Л., Госэнергоиздат, 1963, 552 с. с илл.

6Ф2.13

Редактор *В. В. Енютин* Техн. редактор *Н. И. Борунов*

Слано в набор 16/X 1962 г. Подписано к печати 4/I 1963 г
Бумага 84×108 1/32 28,29 печ. л. Уч.-изд. л. 29,8
Тираж 16 000 экз. Цена 2 р. 24 к. Зак. 2636

Типография Госэнергоиздата. Москва, Шлюзовая наб., 10.

ПРЕДИСЛОВИЕ РЕДАКТОРА ПЕРЕВОДА

О полупроводниках и полупроводниковых приборах издано за последние годы много книг, что неудивительно ввиду практической важности этих приборов для техники и широкого интереса к научным проблемам в этой области.

Однако почти все книги, посвященные этой теме, трактуют специальные вопросы, интересные для конструкторов, технологов или разработчиков полупроводниковых приборов, либо рассчитаны на читателя, не имеющего достаточной физико-математической подготовки. Предлагаемый вниманию читателей перевод книги Дж. Шайва отличается от этих двух категорий книг рядом особенностей. Прежде всего изложение этой книги не рассчитано на специалиста и потому достаточно доступно. Однако оно не элементарно. Популярность изложения книги — результат не упрощения сложных вопросов, а тонкого понимания основ явлений и умелого изложения, рассчитанного на систематическое введение элементарно образованного читателя в круг понятий, идей, физики, конструкций и применения полупроводниковых приборов. В отличие от других книг в этой книге охвачены практически все основные типы полупроводниковых приборов и рассмотрены теоретические основы происходящих в них явлений, включая элементарное изложение основ электронной теории полупроводников.

Наконец, книга эта прямо предназначена для повышения квалификации лиц, имеющих общее и высшее образование, т. е. рассчитана на студентов специальных

технических вузов. Поэтому книга может служить в качестве учебного пособия для студентов вузов и институтов повышения квалификации; в методическом отношении она вполне удовлетворяет этому назначению (за исключением, быть может, последовательности изложения материала).

Сочетание таких важных качеств в одной книге является, на наш взгляд, большим ее достоинством и побудило взяться за ее перевод.

Перевод гл. 1—14 выполнен В. П. Дубенским и И. Г. Морозовой, 15—18 — Т. Г. Медведевой, 19—22— Т. А. Озолиной, 23—25 — Е. К. Херасковой. Некоторые индивидуальности переводчиков, возможно, сказались на общем стиле книги, но редактор стремился максимально выровнять его. В ряде случаев сделаны примечания, исправляющие трактовку автором отдельных вопросов или связанные с некоторым отставанием содержания книги от современного состояния вопроса.

Книга вышла в США в свет в 1959 г., и вследствие этого при бурном развитии полупроводниковой электроники сегодня ее содержание уже не отражает полностью современный уровень этой отрасли электроники. Так, в книге подробно рассматривается точечный триод, тогда как в настоящее время ни один тип точечных триодов более не выпускается промышленностью и его рассмотрение можно считать лишь данью истории.

Появился ряд новых типов полупроводниковых приборов, широко внедрившихся в технику, о которых в книге ничего не могло быть сказано. К ним относится, например, туннельный диод, названный так вследствие того, что переход носителей заряда через тонкую область пространственного заряда p-n-перехода происходит посредством туннельного механизма прохождения сквозь потенциальный барьер. Появился многообещающий триод типа «спейсистор», эмиттер которого находится в слое пространственного заряда коллекторного перехо-

да, что сокращает до минимума время переноса носителей из эмиттера в коллектор и улучшает частотные свойства этого триода. Есть и ряд других типов приборов и методов их применения, о которых читатель сможет найти сведения в дополнительной литературе, указанной переводчиками и редактором в каждой главе. Делать соответствующие добавления или изменения в тексте книги было признано нецелесообразным.

Были сделаны некоторые незначительные сокращения отдельных мест оригинального текста в связи в основном с их рекламным характером.

Наконец, отметим, что обозначения в формулах, единицы измерения величин и их наименования были переведены во всем тексте книги в те, которые приняты в СССР, а терминология изменена по возможности в соответствии с рекомендациями Комитета научно-технической терминологии АН СССР, где сейчас заканчивается разработка терминологии по полупроводникам.

Г. А ТЯГУНОВ

ИЗ ПРЕДИСЛОВИЯ АВТОРА

В 1948 г. фирма Bell Telephone Laboratories создала для вновь принятого технического персонала курсы повышения квалификации с программой, касающейся производства средств связи. В то время фирма Bell System из всех полупроводниковых приборов заинтересовалась почти исключительно термисторами, твердыми выпрямителями и слаботочными диодами. Однако как раз в это время были изобретены транзисторы, и стало ясно, что произошел важный поворот в сторону дальнейшего развития и более широкого использования полупроводниковых приборов. Соответственно этому в новые учебные программы был введен курс полупроводниковых приборов и физики полупроводников.

Впервые этот курс был прочитан в течение 1949—1950 гг. на базе материалов, подготовленных совместно доктором Беккером и мной. С тех пор из года в год содержание курса пополнялось новыми сведениями о разработках полупроводниковых приборов и об исследованиях в области общей физической теории полупроводников и работы полупроводниковых приборов. Текст этой книги представляет собой последний вариант, подготовленный на основе опыта нескольких лет преподавания под постоянным и критическим контролем моих коллег по научно-исследовательскому отделу физики твердого тела и полупроводниковых приборов лаборатории фирмы Bell.

Любое сколько-нибудь полное изучение полупроводников и полупроводниковых приборов должно быть длительным, поскольку эта область уже достаточно развилась и все продолжает расширяться Поскольку рассмотрение только основных вопросов курса полупроводниковых приборов уже требует вдвое большего объема книги, в нашем изложении опущен ряд деталей второстепенной важности. Поэтому читатель напрасно будет искать

здесь упоминания о таких приборах, как гираторы, приборы с изоляторами, ферроэлектрические запоминающие устройства и мазеры (молекулярные усилители), которые, строго говоря, относятся к приборам, использующим электрические явления в твердых телах, но которые тем не менее можно рассматривать как выходящие за пределы более узко ограниченной группы полупроводниковых приборов.

По общему уровню изложения эта книга рассчитана на хорошо подготовленных студентов высших учебных заведений и на лиц с высшим образованием. Особое внимание в ней уделено качественному описанию явлений при различных электронных процессах, происходящих в полупроводниках, так что для читателя, который намерен выбрать полупроводники своей специальностью, эта книга может служить введением к более глубокому физическому и математическому изучению темы. По уровню, глубине и содержанию книга предназначена для заполнения промежутка между элементарным изложением основ физики полупроводников и такими чисто теоретическими работами, как, например, Ф. Зейтц, «Современная теория твердого тела» или В. Шокли, Теория электронных полупроводников.

Материал этой книги логически делится на две основные части: собственно полупроводники и физическую теорию полупроводниковых приборов. Часть первая (гл. 1—14) содержит рассмотрение различных полупроводниковых приборов с эмпирической точки зрения. В этих главах рассмотрены некоторые приборы, их производство, свойства, характеристики и типовые случаи применения, причем изложено достаточное количество начальных сведений, чтобы сделать понятной качественную сторону работы приборов. Во второй части (гл. 15—25) изложена электронная теория твердых полупроводников, чтобы подвести читателя к теоретическому пониманию физической стороны работы полупроводниковых приборов и тем самым к пониманию физической и инженерной сторон разработки полупроводниковых приборов. Таким образом, первая часть книги может быть использована самостоятельно как основа для изучения свойств и поведения полупроводниковых приборов, а вторая часть — как введение в физику полупроводников и конструирование полупроводниковых приборов.

За каждой главой, в которой описан какой-либо полупроводниковый прибор, следует глава о его применении. Такие главы о практическом применении приборов должны облегчить учащимся понимание важности данного прибора и, таким образом, обосновать необходимость его изучения. В этих главах цепи описаны только с функциональной точки зрения, принципиальные схемы упрощены и не содержат технических данных. Описание областей применения приборов не предназначено для замены специальных курсов, целиком посвященных этому вопросу, равно как и содержание этих глав не рассчитано на сообщение сведений, имеющихся в инженерных справочниках, как бы ни были эти сведения необходимы разработчику, намеревающемуся применить прибор в схеме определенного назначения. Рекомендации по практическому применению приборов можно найти в различных технических инструкциях, выпускаемых изготовителями приборов. Кроме того, хотя в разных местах текста изложены некоторые детали физических и химических процессов производства полупроводниковых приборов, эти сведения не могут заменить руководства по технологии изготовления приборов.

Изучение полупроводниковых приборов, как и почти всякого другого вопроса, если оно проводится с воодушевлением и целенаправленностью, не только позволяет узнать нечто полезное, но и служит стимулом для культурного роста. Многие считают, что искусство и наука — это исключающие друг друга области деятельности, находящиеся на противоположных концах шкалы культуры. Мне, однако, кажется, что наука — важная часть культуры и что совместное развитие экспериментальной и теоретической физики, примером которого служит нынешнее состояние знаний о физике твердого тела, представляет собой прекрасный пример современных культурных и интеллектуальных устремлений.

Молодым людям, в руках и умах которых находится будущее науки, многим из них я имел удовольствие и честь быть учителем и другом.

К ЧИТАТЕЛЮ

Если Вы специалист-практик в области электронных приборов или схем или если Вы — ученый, работающий в какой-либо другой области, но желающий пополнить свою общую подготовку или расширить свое знание мира и того, что происходит в нем, то Вы уже достаточно хорошо представляете себе, что заставило Вас читать эту книгу. Вы, несомненно, хорошо отдаете себе отчет о перспективах развития области техники, посвященной полупроводниковым приборам, и Вам излишне указывать на желательность изучения всего нового об этих приборах.

С другой стороны, если Вы — студент, принявшийся за изучение полупроводниковых приборов потому, что они входят в программу, или потому, что Ваш руководитель посоветовал ознакомиться с ними, то Вы можете призадуматься, стоит ли тратить время на изучение предмета, который кажется Вам относительно второстепенным.

Подобные сомнения связаны с недооценкой важности полупроводниковых приборов. На самом-то деле начинающий инженер или студент не сможет сколько-нибудь успешно составить практическую схему для промышленной или связной аппаратуры, если он будет избегать применения полупроводников.

Автор стремился описать различные полупроводниковые приборы, обсудить их свойства и разъяснить ме-

ханизм их работы с помощью основных физических понятий, которые вводятся по мере надобности. Глубина изложения материала такова, что изучив эту книгу, читатель сможет следить по технической литературе за дальнейшим развитием этой области техники.

Изложение начинается с очень легкого обзора полупроводниковых материалов и их свойств Этот обзор дает общее знакомство с материалом, который далее рассмотрен с большими математическими подробностями Далее описаны сами полупроводниковые приборы Для каждого из приборов указано, что́ он собой представляет, как изготовляется, как работает и для чего применяется Затем свойства полупроводников разобраны в деталях, исходя из основ физики Показано, чем обусловлены уникальные свойства их

Зная теперь наши цели и пути, которыми мы придем к их осуществлению, примемся за само изучение. И я пожелаю всем Вам, читателям этой книги, того волнения, которое я испытывал и продолжаю испытывать, изучая полупроводниковые приборы и как и почему они работают

Автор.

СОДЕРЖАНИЕ

Предисловие редактора перевода 3
Из предисловия автора 6
К читателю 9

Часть первая

ПРИРОДА, СВОЙСТВА, ПОВЕДЕНИЕ И ПРИМЕНЕНИЕ ПОЛУПРОВОДНИКОВЫХ ПРИБОРОВ

Глава первая. **Введение** 17
Литература 22

Глава вторая. **Предварительное ознакомление с природой и свойствами полупроводников** 22
2-1. Определения и примеры полупроводников 22
2-2. Удельное сопротивление и температурный коэффициент 29
2-3. Выпрямительные свойства полупроводников 36
2-4. Фотоэлектрические свойства полупроводников 38
2-5. Двуполярная проводимость в полупроводниках 40
2-6. Термоэлектрические явления в полупроводниках 42
2-7. Эффект Холла в полупроводниках 43
Задачи 45
Литература 46

Глава третья. **Термосопротивления (термисторы)** . . 46
3-1. Общие сведения 46
3-2. Изготовление термисторов прямого подогрева 48
3-3. Физические свойства самонагревающихся термисторов 51
3-4. Подогревные термисторы 70
3-5. Старение и шумы 71
3-6. Замечания о выборе режима и графическом расчете цепей 74
Задачи 76
Литература 77

Глава четвертая **Применение термисторов** 77
4-1. Термометрия 78
4-2. Болометрия с помощью термисторов 80
4-3. Температурная компенсация 82
4-4. Измерение высокочастотной мощности 84

4-5. Управление характеристиками электронных цепей . . 85
4-6. Стабилизация напряжения 88
4-7. Измерение давления 89
4-8. Выдержка времени и схемы задержки 91
4-9. Системы пожарной сигнализации 93
4-10. Усилители с отрицательным сопротивлением 95
Задачи 96
Литература 96

Глава пятая. **Варисторы. Введение и описание плоскостных приборов** 97

5-1. Несимметричные варисторы 98
5-2. Симметричные варисторы—тирит и карбид кремния . . 119
Задачи 122
Литература 122

Глава шестая. **Свойства варисторов; поведение их при высокой частоте; точечно-контактные приборы** 123

6-1. Сопротивление постоянному и переменному току . . . 123
6-2. Барьерная емкость 126
6-3. Полное сопротивление выпрямителя и его поведение при малых переменных сигналах 129
6-4. Поведение выпрямителей при больших переменных сигналах 134
6-5. Выпрямители с распределенным сопротивлением и точечно-контактные 137
Задачи 148
Литература 149

Глава седьмая. **Применение варисторов** 149

7-1. Устройство для заряда аккумуляторов 149
7-2. Силовые выпрямители 151
7-3. Модуляторы и демодуляторы 153
7-4. Экспандеры и компрессоры 159
7-5. Устройства защиты от перегрузок, подавители коммутационных помех, ограничители напряжения и схемы молниезащиты 162
7-6. Устройства для защиты контактов реле 164
7-7. Ограничители амплитуды сигналов переменного тока . 165
7-8. Диодные логические схемы 167
Задачи 168
Литература 170

Глава восьмая. **Полупроводниковые фотоэлементы** 171

8-1. Фотосопротивления 172
8-2. Фотогальванические элементы 189
8-3. Прочие фотоэлементы различных типов 196
8-4. Заключительные замечания 197
Задачи 202
Литература 203

Глава девятая. **Применение полупроводниковых фотоэлементов** 205

9-1. Цепи постоянного тока 205
9-2. Использование фотоэлементов в качестве управляемых светом ключей при высоких уровнях сигналов 206
9-3. Цепи переменного тока 213
9-4. Генерирование энергии 214
Задачи 216
Литература 216

Глава десятая. **Транзисторы** 217

10-1. Введение 217
10-2. Точечные транзисторы 219
10-3. Характеристики точечных транзисторов 223
10-4 Элементарная теория механизма действия транзистора 228
10-5. Плоскостные транзисторы 231
Задачи 239
Литература 240

Глава одиннадцатая. **Характеристики транзисторов и элементы схем** 241

11-1. Графические характеристики 242
11-2. Описания с помощью эквивалентных схем 245
11-3. Транзистор как четырехполюсник 250
11-4. Взаимосвязь между тремя способами описания транзисторов 254
11-5. Цепи транзисторов 257
11-6. Схемы с общим эмиттером и общим коллектором . . 259
Задачи 265
Литература 265

Глава двенадцатая. **Работа транзисторов на высоких частотах и высокочастотные транзисторы** . . . 266

12-1. Зависимость α от частоты 267
12-2. Влияние емкостей переходов на частоту 271
12-3. Плоскостные тетроды 275
12-4. Поверхностно-барьерные триоды 277
12-5. Диффузионные транзисторы 279
12-6. Транзисторы с коллекторной нейтральной прослойкой *p-n-i-p* или *n-p-i-n* 283
Литература 285

Глава тринадцатая. **Прочие типы транзисторов** . . 285

13-1. Стержневой транзистор с модулированной проводимостью 285
13-2. Канальные транзисторы 288
13-3. Лавинные транзисторы 290

13-4. Ловушечные триоды и триоды типа *p-n-p-n* 293
13-5. Силовые транзисторы 296
Задачи 298
Литература 299

Глава четырнадцатая. **Применение транзисторов** 299

14-1. Усилители 300
14-2. Генераторы и мультивибраторы 303
14-3. Импульсные схемы с транзисторами 305
14-4. Селекторные и ключевые схемы с транзисторами . . . 307
Задача 311
Литература 311

Часть вторая

ФИЗИКА ТВЕРДЫХ ТЕЛ И ЕЕ ПРИМЕНЕНИЕ К КОНСТРУИРОВАНИЮ ПОЛУПРОВОДНИКОВЫХ ПРИБОРОВ

Глава пятнадцатая. **Зонная теория твердых тел** . . 312

15-1. Зонная теория с точки зрения атомистического приближения 313
15-2. Зонная теория с точки зрения коллективного электронного приближения 328
15-3. Итоги 340
Задачи 341
Литература 341

Глава шестнадцатая. **Механизмы проводимости твердых тел** 341

16-1. Собственная проводимость полупроводников 343
16-2. Примесная проводимость в полупроводниках; доноры и акцепторы 347
16-3. Подвижность и удельная электропроводность 353
Задачи 362
Литература 363

Глава семнадцатая. **Статистика электронов и дырок** 364

17-1. Введение 364
17-2. Статистика Ферми 366
17-3. Случай собственного полупроводника 368
17-4. Применение теории к полупроводнику *p*-типа 375
17-5. Применение теории к полупроводнику *n*-типа 382
17-6. Применение теории к полупроводнику, содержащему как доноры, так и акцепторы 385
17-7. Постоянство произведения *pn* 389
Задачи 390
Литература 391

Глава восемнадцатая. **Уравнение непрерывности для концентраций избыточных неосновных носителей** 391

18-1. Общие замечания относительно уравнений непрерывности 391
18-2. Уравнение непрерывности для концентраций избыточных неосновных частиц в полупроводниках 392
18-3. Решения уравнения непрерывности для частных случаев 396
18-4. Соотношения Эйнштейна 401

Задачи 404

Литература 405

Глава девятнадцатая. **Теория *p-n* перехода** 405

19-1. *p-n* переход в равновесном состоянии 405
19-2. Обратный и прямой токи через *p-n* переходы 416
19-3. Использование *p-n* переходов в качестве эмиттеров . 422
19-4. Емкость барьера 427

Задачи 430

Литература 431

Глава двадцатая. **Теория контактов металл—полупроводник** 431

20-1. Расчет электрической характеристики идеального выпрямителя 431
20-2. Контакт металла с полупроводником *p*-типа 437
20-3. Выпрямительные ряды 441
20-4. Влияние поля электрического изображения на характеристику перехода 445
20-5. Точечные контакты как эмиттеры неосновных носителей 450

Задачи 451

Литература 451

Глава двадцать первая **Теория плоскостного триода** 451

21-1. Прямой ток эмиттера и эффективность эмиссии неосновных носителей 454
21-2. Входное полное сопротивление эмиттера 457
21-3. Эффективность переноса частиц от эмиттера к коллектору 459
21-4. Усиление тока в коллекторном переходе 461
21-5. Величина α и определяющие ее множители 464
21-6. Коллекторный ток насыщения 466

Задача 467

Литература 467

Глава двадцать вторая. **Оптические свойства полупроводников и фотоэлектрические эффекты** . . 468

22-1. О квантовой природе света 468
22-2. Фотопроводимость 476
22-3. Фотогальванические эффекты 486
22-4. Фотоэлектронные умножители 488
22-5. Флуоресценция и фосфоресценция в полупроводниках 491
22-6. Фотоэлектронная эмиссия полупроводников 494

Задачи 499

Литература 499

Глава двадцать третья. **Термоэлектрические эффекты в полупроводниках** 500

23-1. Вычисление термо-э.д.с. и удельной термо-э.д.с. . . . 504
23-2. Эффекты Пельтье и Томсона 506
23-3. Соотношение между термо-э.д.с. и коэффициентами π и θ 510
23-4. Другие полезные термоэлектрические соотношения . . 513
23-5. Сравнение с экспериментальными данными 517

Задачи 518

Литература 519

Глава двадцать четвертая. **Эффект Холла** 519

24-1. Происхождение напряжения Холла 519
24-2. Измерение величин напряжения и константы Холла . 524
24-3. Результаты 527

Задачи 531

Литература 531

Глава двадцать пятая. **Некоторые вопросы технологии изготовления приборов** 531

25-1. Приготовление и очистка германия 532
25-2. Создание переходов методом вытягивания кристаллов 539
25-3. Выращивание с переменной скоростью 543
25-4. Сплавление 545
25-5. Диффузионные переходы 548
25-6. Общие замечания 550

Задачи 551

Литература 551

ЧАСТЬ ПЕРВАЯ

ПРИРОДА, СВОЙСТВА, ПОВЕДЕНИЕ И ПРИМЕНЕНИЕ ПОЛУПРОВОДНИКОВЫХ ПРИБОРОВ

ГЛАВА ПЕРВАЯ

ВВЕДЕНИЕ

По сравнению с хорошо развитой областью вакуумной электроники многие считают полупроводниковую электронику не вполне полноценной отраслью науки. Лишь в последние годы полупроводники и полупроводниковые приборы вошли в арсенал повседневных средств разработчиков цепей и целых систем. Однако полупроводниковые приборы имеют более древнюю историю, чем вакуумные приборы, и появились раньше.

Некоторые свойства, которые мы теперь связываем с полупроводниками, были известны более чем за столетие до нас. В 1833 г. Фарадей [Л. 1] указал, что сернистое серебро обнаруживает отрицательный температурный коэффициент сопротивления. В. Смит [Л. 2] в 1873 г. обнаружил, что сопротивление элемента схемы, сделанного из кристаллического селена, понижается при его освещении. Одну из первых описанных в литературе попыток использовать свойства полупроводников сделал около трех четвертей столетия назад Александр Белл [Л. 3], разработавший световой клапан, который модулировал поток света звуковой частотой. Из этого модулированного потока коллиматором выделялся параллельный луч, который проходил некоторое расстояние, после чего он улавливался конденсором и концентрировался на селеновом приемном элементе; электрический сигнал на выходе этого элемента являлся воспроизведением входного сигнала передающего устройства. «Фотофон» Белла в известной степени можно считать предшественником современных радиорелейных линий, в которых информация передается путем модуляции направленного излучения.

О выпрямительном эффекте в контактах между полупроводниками и металлами впервые сообщил Ф. Браун [Л. 4] в 1874 г. Однако это открытие принесло мало пользы, пока в первом десятилетии нашего века не были разработаны практические типы детекторов для радиоприема [Л. 5, 6 и 7]. После короткой вспышки популярности кристаллические детекторы были вытеснены вакуумными диодами, которые господствовали в течение второго десятилетия. Только в 30-х годах заинтересованность разработчиков в освоении все более и более высоких частот, недоступных с помощью обычных электронных ламп, вновь вызвала к жизни кристаллические детекторы [1]. В течение второй мировой войны кристаллические детекторы были единственными приборами, пригодными для детектирования и демодуляции на частотах радиолокационного диапазона. Однако в наше время в этот частотный диапазон проложили себе дорогу лампы бегущей волны, и никто не может предсказать, какой прибор будет следующим этапом в этой чередующейся последовательности.

В конце 20-х — середине 30-х годов появились медно-закисные [Л. 8] и селеновые [Л. 9] мощные вентили. Вскоре они нашли применение в устройствах для заряд-ки аккумуляторов, в выпрямителях для питания электронных устройств, в модуляторах и демодуляторах для систем дальней связи с передачей на несущей частоте. В конце 30-х — начале 40-х годов были разработаны термисторы, усовершенствованы и специализированы диоды и фотоэлементы. Последний член семьи полупроводников — транзистор — появился в 1948 г. Его изобретение и продолжающееся усовершенствование дали колоссальный толчок широкому использованию полупроводниковых приборов всех видов.

Интересно, что эта картина развития полупроводниковых приборов очень похожа на развитие достижений в других отраслях науки, постепенно приводящих к появлению практически полезных результатов, будь то

---

[1] Автор не упоминает о работах сотрудника Нижегородской радиолаборатории О. В. Лосева в 1921—1925 гг. над полупроводниковыми приборами О. В. Лосев разработал первый полупроводниковый прибор с отрицательным сопротивлением («кристадин»), применявшийся в качестве усилителя и генератора электрических колебаний на частотах вплоть до 12 *Мгц. Прим. ред.*

кремнийорганические смолы, приборы для тугоухих или турбины для атомных электростанций. Вначале появляются разрозненные наблюдения в данной области. Затем делается несколько экспериментальных попыток применить эти наблюдения для практических целей. Число этих попыток все увеличивается, и они следуют все чаще по мере опубликования новых эмпирических данных. Разрабатываются и поступают в продажу примитивные производственные образцы приборов. Однако по настоящему целенаправленное, согласованное и широкое развитие новой области начинается не раньше, чем возникнут одновременно три условия 1) достаточно большая потребность, чтобы оправдать это расширение; 2) развитие теоретического понимания, достаточного для объяснения имеющихся эмпирических данных и ведения дальнейших разработок, 3) создание соответствующей технологии, обеспечивающей возможность точно контролируемого, воспроизводимого приготовления материалов и изготовления приборов.

Существующее многообразие видов полупроводниковых приборов показано на рис. 1-1 На этой схеме полупроводниковые приборы подразделены вначале на классы, указанные в кружках. Более мелкие подразделения по видам, строению и технологии производства показаны в прямоугольниках, соединенных с кружками. Такая схема дает наглядное представление о сложности всего семейства полупроводниковых приборов. Однако эта схема не дает указаний о степени использования различных приборов в разных других отраслях промышленности и связи, так же как и сведений о количестве типов приборов в отдельных группах. Один прямоугольник на схеме может изображать десятки типов приборов, отличающихся друг от друга только величиной какого-либо параметра.

Выполнение с помощью полупроводниковых приборов таких элементарных функций, как генерирование колебаний, усиление, модуляция и демодуляция, выпрямление, генерирование импульсов, переключение, стало сейчас повседневным явлением. Используя различные полупроводниковые приборы совместно с пассивными элементами цепей для выполнения этих элементарных функций в различных сочетаниях, разработчики создали основные системы функциональных блоков, такие,

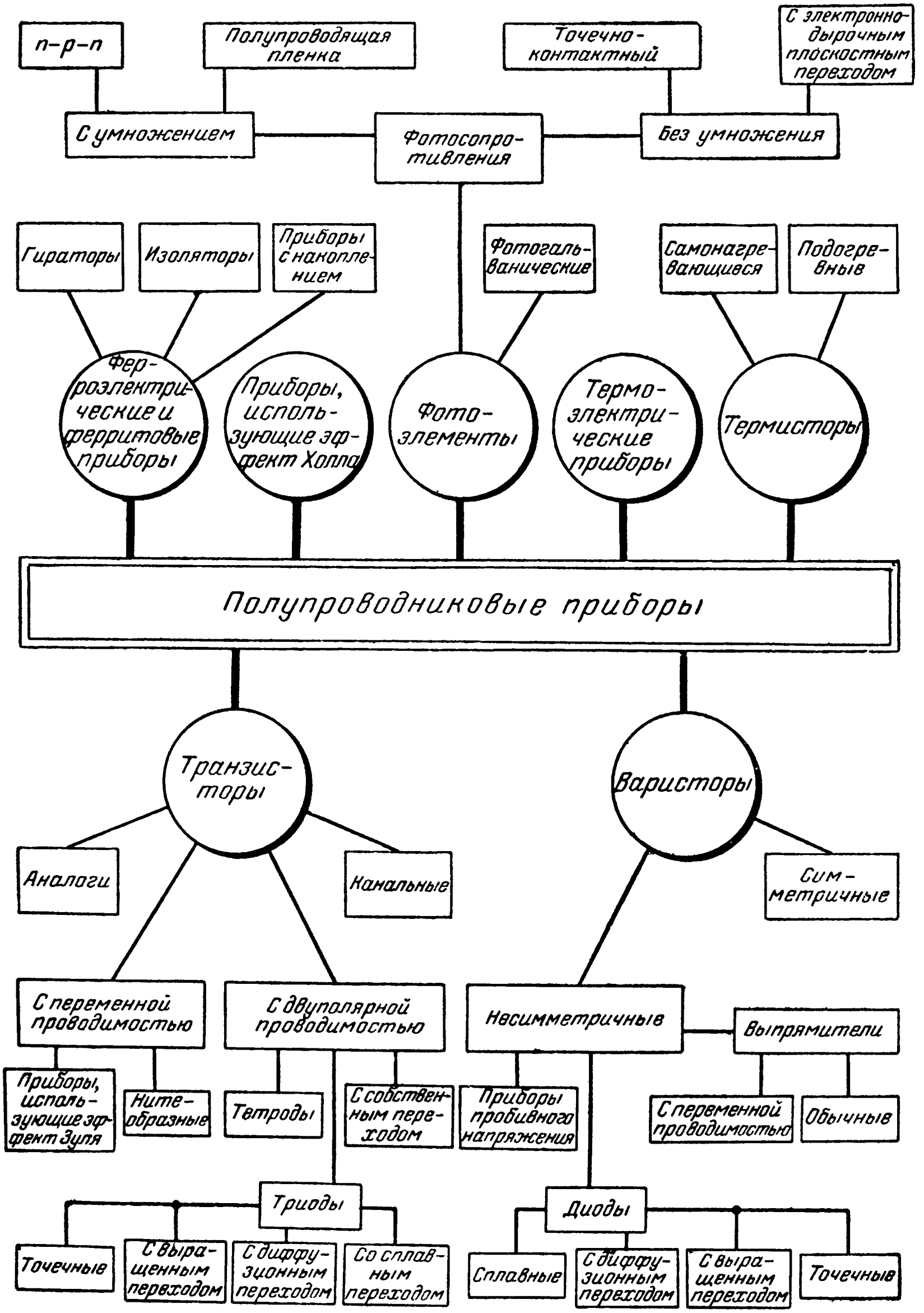


Рис. 1-1. Классификация полупроводниковых приборов.

как двоичные вычислительные устройства, блоки накопления информации для вычислительных устройств, кодирующие и декодирующие приборы, системы уплотнения линий связи, промежуточные линейные усилители, радио- и телевизионные приемники, трансляционные установки и т. д. Эти функции обычно выполняются полу-

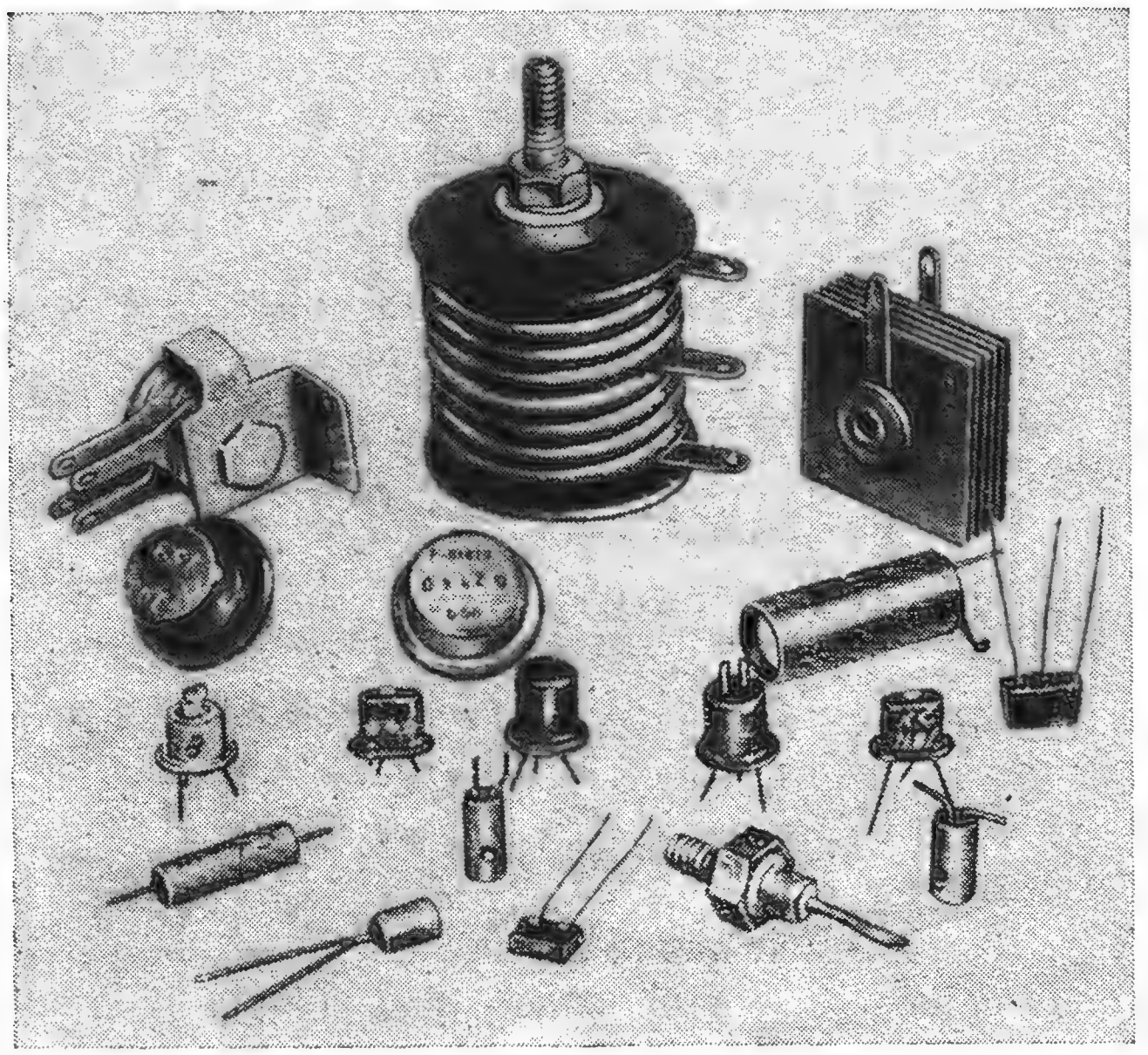


Рис 1-2 Внешний вид некоторых полупроводниковых приборов

проводниковыми приборами с большей экономией стоимости, места и потребляемой мощности, чем другими электронными устройствами, например электронными лампами и реле. При нынешней тенденции к усложнению электронной аппаратуры «миниатюризация» из удобства становится необходимостью. Полупроводниковые приборы не только сами уменьшились в габаритах, но и стимулировали разработку миниатюрных трансформаторов, конденсаторов и сопротивлений.

На рис 1-2 показан внешний вид некоторых полупроводниковых приборов.

ЛИТЕРАТУРА

1. Faraday Michael, Experimental Researches in Electricity, Bernard Quaritch, London, 1839, vol. 1, p. 122.
Русский перевод: Фарадей М., Экспериментальные исследования по электричеству, Изд. АН СССР, 1947.
2. Smith, Willoughby, The Action of Light on Selenium, Journal of the Society of Telegraph Engineers, 1873, vol. 2, p. 31.
3. «Bell's Photophone», Nature, 1880, vol. 22, p. 15.
4. Braun F., Über die Stromleitung durch Schwefelmetalle, Annalen der Physik und Chemie, 1874, Bd. 153, S. 556.
5. Bose J. C., U. S. Patent No. 755840, 1904.
6. Dunwoody H. H. C., U. S. Patent No. 837616, 1906.
7. Pierce G. W., Crystal Rectifiers for Electric Currents and Electrical Oscillators, Physical Review, 1907, vol. 25, p. 31; 1909, vol. 28, p. 153.
8. Grondahl L. O. and Geiger P. H., A New Electronic Rectifier, Transactions of the American Institute of Electrical Engineers, 1927, vol. 46, p. 357.
9. Lange B., Die Photoelemente und ihre Anwendung, Johann Ambrosius Barth, Leipzig, 1936.

ГЛАВА ВТОРАЯ

ПРЕДВАРИТЕЛЬНОЕ ОЗНАКОМЛЕНИЕ С ПРИРОДОЙ И СВОЙСТВАМИ ПОЛУПРОВОДНИКОВ

В этой главе мы познакомимся с некоторыми наиболее существенными свойствами полупроводников. В частности, рассмотрим, как эти свойства отличаются в качественном или количественном отношении от свойств «обычных» веществ, и покажем, как особенности полупроводников могут быть использованы для создания полупроводниковых приборов. Вначале мы затронем лишь качественную сторону поведения полупроводников, изложив теоретические основы, объясняющие это поведение, в последующих главах.

2-1. ОПРЕДЕЛЕНИЯ И ПРИМЕРЫ ПОЛУПРОВОДНИКОВ

Возможно, ни одно физическое свойство веществ в природе не характеризуется таким огромным диапазоном значений от наименьших до наибольших, как электрическое удельное сопротивление. Начиная с наиболее проводящих металлов и кончая самыми лучшими диэлек-

триками, этот диапазон включает по крайней мере 25 порядков. Фактически эта величина еще выше, но наша техника измерений не в состоянии охватить ее крайние значения. На рис 2-1 приведена шкала, на которой указаны величины удельных сопротивлений некоторых широко распространенных веществ при комнатной температуре. Удельные сопротивления веществ, определяемых как «хорошие проводники» (металлы), находятся в пре-

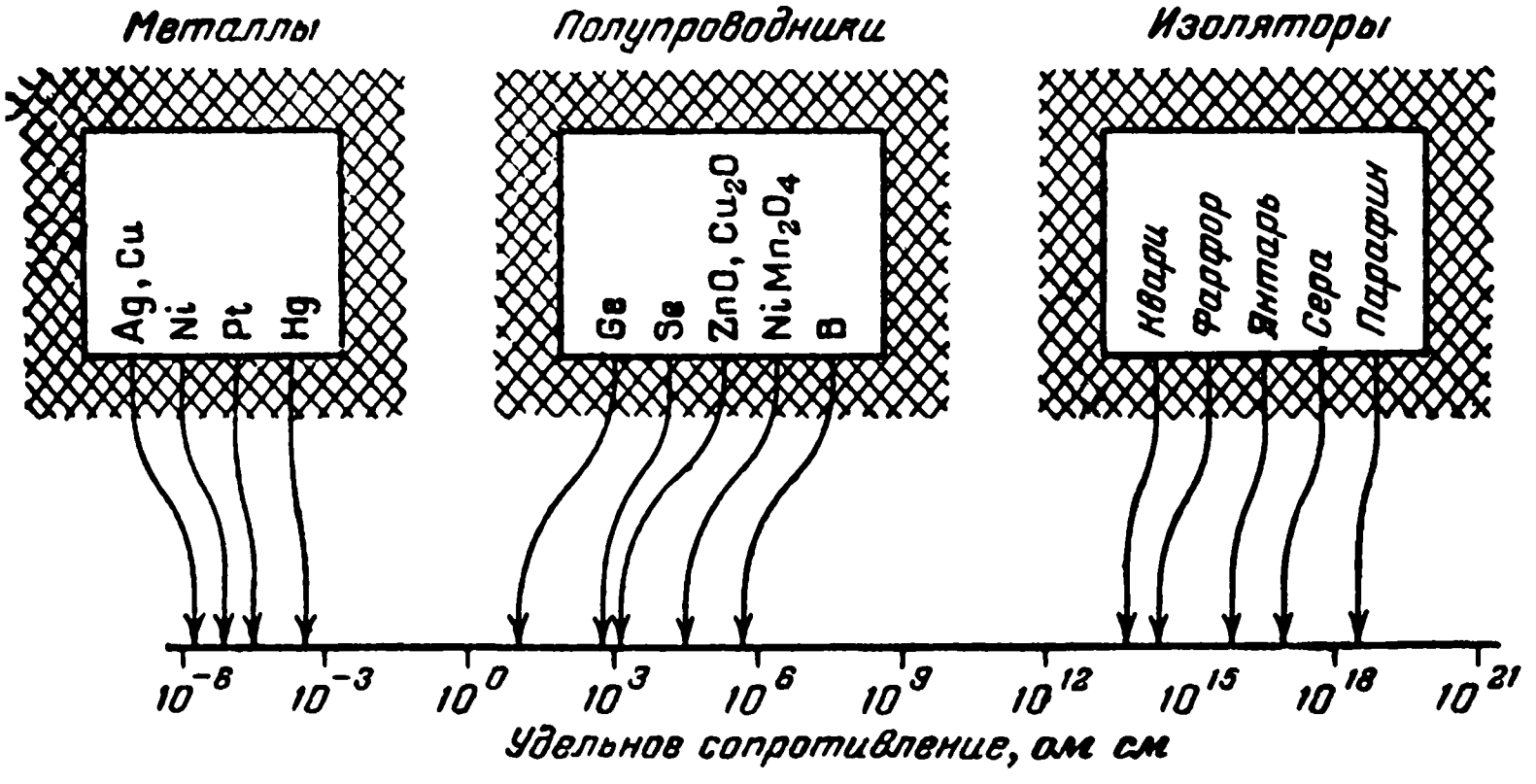


Рис 2-1 Удельные сопротивления некоторых хорошо известных веществ при комнатной температуре

делах от 10^{-6} до приблизительно 10^{-3} *ом·см* *, тогда как вещества, определяемые как изоляторы, имеют удельное сопротивление выше приблизительно 10^{12} *ом·см*. Между металлами и изоляторами на шкале показан ряд веществ, обычно мало известных. Они слишком плохо проводят электричество, чтобы их можно было использовать для соединительных проводов и кабелей, и в то же время создают слишком большие утечки, препятствующие их использованию для линейных изоляторов и установочной арматуры. Такие промежуточные вещества называются полупроводниками Диапазон ве-

---

* В настоящем обзоре использованы на более или менее равных правах две физические величины удельное сопротивление, измеряемое в *ом·см*, и величина, обратная ему — удельная проводимость, измеряемая в 1/*ом·см*, или *мо/см*. В необходимых случаях в зависимости от удобства, например при рассмотрении полного сопротивления или полной проводимости цепей, будет применяться та или иная величина

личин удельного сопротивления, в пределах которого данное вещество следует отнести к полупроводникам, точно не определен, и никто не делал таких попыток. Обычно вещества, имеющие удельные сопротивления в пределах от 10^{-3} до 10^{10} *ом·см,* относятся к полупроводникам, они обнаруживают типичные для полупроводников физические свойства.

С первого взгляда полупроводники могут показаться сравнительно бесполезными материалами. Однако им присущи не наблюдающиеся в заметной степени в металлах и изоляторах свойства, привлекающие к полупроводникам самое пристальное внимание. Из этих материалов изготовляются полупроводниковые диоды, триоды, термисторы и фотоэлементы, которые ныне проникли в электронику и революционизировали ее.

Среди полупроводников можно встретить некоторые простые химические вещества, такие, как кремний, германий, селен и бор. Многие смеси соединений металлов с кислородом, серой, селеном, теллуром, галоидами и углеродом также являются полупроводниками. Вдобавок имеются широкие возможности для создания более сложных материалов путем смешивания окислов различных металлов и сплавления их в таких пропорциях, чтобы получить некоторые определенные типы кристаллических структур. Всего изучено несколько сотен веществ, отнесенных к полупроводникам. В табл. 2-1 перечислены лишь некоторые наиболее известные вещества.

Полупроводники можно классифицировать также в соответствии с природой частиц, являющихся носителями заряда в них. Так, существуют ионные и электронные полупроводники. В ионных полупроводниках ток создается перемещением ионов того вещества, из которого состоит полупроводник. Примером полупроводника с ионной проводимостью является хлористое серебро. Ток в этом веществе создается главным образом за счет движения ионов серебра, которые покидают свои места в правильной кристаллической решетке и перемещаются, как в электролитах, сквозь кристалл под действием электрического поля. Такой процесс сопровождается переносом вещества, образованием поляризованных слоев и отложением вещества на электродах с последующими местными изменениями химического состава вещества. Таким образом, полупроводники с ионной проводимо-

Таблица 2-1

Некоторые полупроводниковые материалы

Химическая формула	Название	Класс химических соединений
Ge	Германий	Простые вещества
Si	Кремний	
Se	Селен	
Te	Теллур	
B	Бор	
Cu_2O	Закись меди	Окислы металлов
ZnO	Окись цинка	
NiO	Закись никеля	
Mn_2O_3	Окись марганца	
PbO_2	Двуокись свинца	
PbS	Сульфид свинца	Сульфиды металлов
Tl_2S	Сульфид таллия	
$FeFe_2O_4$	Закись—окись железа	Смешанные окислы типа шпинеля
$ZnFe_2O_4$	Феррит цинка	
$NiMn_2O_4$	Манганит никеля	
$MgTi_2O_4$	Титанит магния	
ZnSe	Селенид цинка	Различные соединения
Ag_2Te	Теллурид серебра	
CuJ	Йодид меди	
B_4C_3	Карбид бора	
SiC	Карбид кремния	

стью нестабильны по своим свойствам и не используются для производства полупроводниковых приборов. Они упомянуты здесь лишь для полноты изложения, и мы не будем возвращаться к ним.

В электронных полупроводниках носителями электричества являются электроны. Само вещество с электронной проводимостью при прохождении через него электрического тока не изменяется, так как взамен электронов, уходящих из рассматриваемого полупроводника к положительному электроду, из отрицательного электрода поступают новые электроны подобно протеканию воды через трубу.

Электронные полупроводники подразделяются дальше на подклассы в зависимости от того, чем в основном создается в них ток — отрицательными электронами

или положительными «дырками». Положительная дырка представляет собой место в электронной структуре твердого тела, где может находиться электрон, но где его нет. Поэтому дырка ведет себя во всех отношениях, как частица, имеющая заряд, массу и подвижность, как у настоящего электрона, но знак заряда обратный, т. е. положительный. Положительная дырка в некоторых отношениях может быть сравнена с пузырьком в жидкости, который представляет собой место, где должна быть жидкость, но где ее нет. В то время как сила тяжести на все тела действует вниз, пузырек стремится двигаться вверх, как если бы сила тяжести действовала на него в обратном направлении. Подобным же образом дырка в электрическом поле движется в сторону, противоположную движению обычного свободного электрона. Таким образом, положительная дырка «реальна» лишь в такой же степени, как и пузырек в жидкости. Однако представление о положительных дырках позволяет достаточно полно понять некоторые особенности поведения полупроводников; иначе было бы весьма трудно представить их в виде какой-либо физической модели. В следующих главах мы рассмотрим многие такие особенности, а далее, в гл. 15, 16 и 17, мы сможем яснее установить, что же представляет собой дырка, и понять ее поведение. В настоящий момент мы допустим существование дырок наравне с электронами и будем рассматривать их как один из видов элементарных частиц в природе.

Если электропроводность полупроводника обусловливается преобладанием в нем свободных электронов, то такой полупроводник относится к *n*-типу (начальная буква слова negative, т. е. отрицательный, относящегося к полярности частиц, переносящих в данном случае электричество). Если в полупроводнике преобладают положительные дырки, то он относится к *p*-типу (от слова positive, т. е. положительный). Роль электронов и дырок в передаче электрического тока через полупроводники *n*-типа и *p*-типа поясняется с помощью двух схем, приведенных на рис. 2-2. В полупроводнике *n*-типа ток создается с помощью отрицательных электронов, которые перемещаются сквозь полупроводник от отрицательного к положительному его концу и далее следуют через внешнюю цепь подобно жидкости в замкнутом

трубопроводе, причем роль насоса выполняет источник э. д. с., например батарея. В цепи, содержащей полупроводник *p*-типа, ток в участках из металлических проводов создается, как и в предыдущем случае, потоком электронов. Однако внутри полупроводника электричество переносится положительными дырками, движущимися в обратном направлении — от положительного к отрицательному электроду полупроводника.

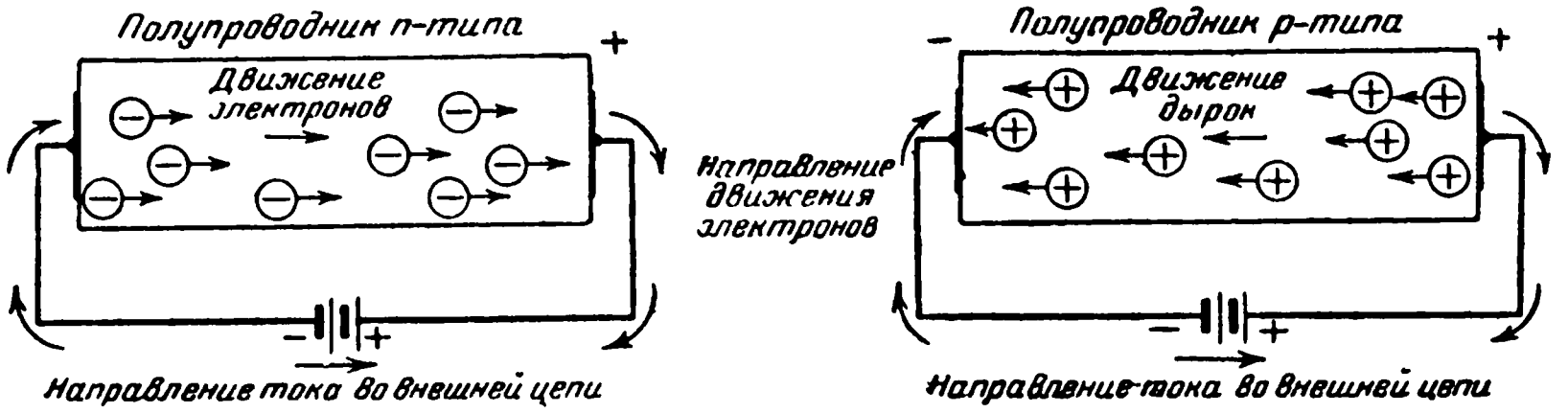


Рис 2-2 В полупорводнике *n*-типа (слева) электроны движутся от минуса к плюсу; в полупроводнике *p*-типа (справа) дырки движутся от плюса к минусу В обоих случаях направление тока одинаково.

Чтобы яснее представить себе механизм проводимости *p*-типа, проследим движение электронного заряда через цепь, включающую полупроводник *p*-типа (изображенную на рис. 2-2), и посмотрим, что происходит с ним, начиная с правого контактного электрода на полупроводнике. Здесь электрон извлекается во внешнюю цепь из электронного газа внутри полупроводника, оставляя в полупроводнике положительную дырку. Электрон проходит через соединительные провода, батарею и попадает в левый электрод. Тем временем дырка движется внутри полупроводника справа налево. На левом электроде происходит рекомбинация электронов и дырок из числа миллионов частиц, прошедших указанный цикл. В результате во всех участках цепи идет одинаковый ток. В течение всего времени прохождения тока общее число дырок в полупроводнике остается постоянным. На каждую дырку, исчезающую на левом электроде за счет рекомбинации, возникает новая дырка на правом электроде, и этот процесс продолжается, пока идет ток.

На рис. 2-2 показано, что направление тока во внешней цепи одинаково для полупроводников *n*-типа и *p*-типа. Ток, создаваемый отрицательными электронами,

движущимися в каком-либо направлении, направлен в ту же сторону, что и ток, создаваемый положительными дырками, движущимися в обратном направлении. Ни в одной из этих двух цепей нет никаких признаков, на основании которых можно было бы экспериментально установить принадлежность полупроводника к тому или иному типу проводимости. Однако имеются другие способы измерений, с помощью которых можно установить тип проводимости полупроводника. Например, полупроводники *n*-типа и *p*-типа обладают односторонней проводимостью противоположных направлений, если их использовать в выпрямительном контакте с соответствующим металлом. Далее, если изготовить термопару, в которой в качестве одного из электродов использован полупроводник, то полярность образующейся э. д. с. будет зависеть от типа проводимости взятого полупроводника. Полярность э. д. с. фотоэлемента, изготовленного из полупроводника, также определяется типом проводимости полупроводника. Эффект Холла, наблюдающийся при пропускании электрического тока через полупроводник, помещенный в магнитное поле, заключается в возникновении напряжения в направлении, поперечном относительно тока, причем полярность этого напряжения зависит опять-таки от типа проводимости полупроводника. В дальнейшем мы вернемся к более детальному рассмотрению этих явлений, поскольку их понимание совершенно необходимо для уяснения механизма работы полупроводниковых приборов.

Некоторые полупроводники обнаруживают только один тип проводимости. Так, закись меди и селен всегда имеют проводимость *p*-типа, а окись цинка — *n*-типа. Кремнию и германию по желанию может быть придана проводимость любого типа путем введения соответствующих примесей. Действительно, в различных участках одного и того же образца полупроводника можно наблюдать проводимости различного типа. Это явление подсказывает, что в производстве полупроводников степень чистоты исходных материалов играет существенную роль. Изучение методов контроля, природы, количества и распределения примесей в полупроводниковых материалах при производстве полупроводниковых приборов так же важно, как и понимание основных физических процессов в полупроводниках.

Если бы было возможно изготовить полупроводник достаточно высокой степени чистоты, то он не обнаружил бы какого-либо определенного типа проводимости (*n* или *p*). Его поведение было бы в какой-то степени промежуточным; такой тип проводимости принято называть *собственной* проводимостью полупроводника. Однако чтобы изготовить образец германия, который имел бы присущую этому веществу собственную проводимость при комнатной температуре, необходимо, чтобы содержание примесей в нем не превышало одного атома на 100 млн. атомов основного вещества, а для изготовления кремния с собственной проводимостью загрязненность должна быть не выше одного атома на 10 триллионов атомов основного вещества. Такая степень чистоты выше возможностей современной технологии, и никто еще не сообщил об изготовлении кремния с собственной проводимостью, хотя германий с собственной или почти собственной проводимостью изготовлялся неоднократно. В последующих главах подробно разобран смысл терминов «проводимость *n*-типа», «проводимость *p*-типа» и «собственная проводимость» в свете основ физики твердых тел, а также применительно к разработке полупроводниковых приборов.

2-2. УДЕЛЬНОЕ СОПРОТИВЛЕНИЕ И ТЕМПЕРАТУРНЫЙ КОЭФФИЦИЕНТ

Удельное сопротивление полупроводника в значительной степени зависит от температуры. На рис. 2-3 показаны величины удельных сопротивлений некоторых типичных полупроводников, представленные в логарифмическом масштабе как функции температуры. Логарифмический масштаб по оси ординат необходим для того, чтобы охватить огромный диапазон величин удельных сопротивлений, наблюдаемых у полупроводников в указанном интервале температур. На том же графике для сравнения приведена соответствующая кривая для платины — типичного металла. По форме кривых удельных сопротивлений видно, что полупроводники имеют температурный коэффициент *отрицательного* знака (сопротивление понижается при повышении температуры) и значительно большей величины, чем у металлов. Удельное сопротивление типичного металла повышается вдвое при повышении температуры примерно на 300° С

относительно комнатной температуры. Что же касается полупроводников, то можно изготовить материал, удельное сопротивление которого понижается вдвое при повышении температуры всего на 18° С.

Ясно, что такую температурную зависимость можно использовать для практических целей. Например, полу-

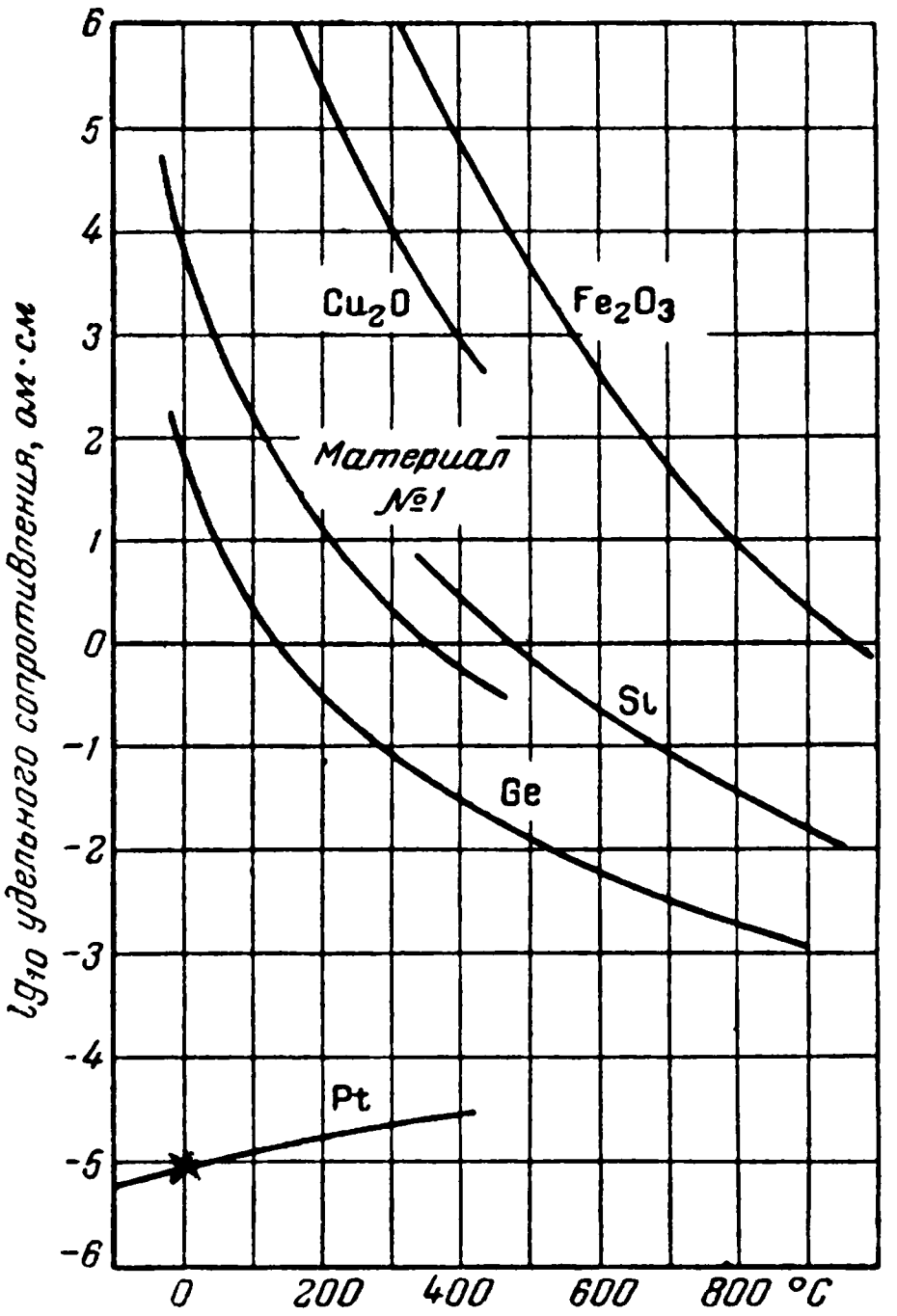


Рис 2-3. Зависимости удельного сопротивления типичных полупроводников и металлов от температуры

проводниковый термометр сопротивления должен быть более чувствительным, чем сделанный из металла. Полупроводниковые приборы, в которых используется температурная зависимость свойств полупроводников, называются термисторами или температурно зависимыми сопротивлениями. Эти приборы описаны в гл. 3.

Встречаясь с кривой, подобной одной из показанных на рис. 2-3, инженер обычно пробует различные способы

построения графиков, чтобы определить, можно ли эту зависимость представить в виде прямой линии.

Кривые на рис. 2-3 могут быть преобразованы в почти прямые линии, если откладывать по оси ординат ло-

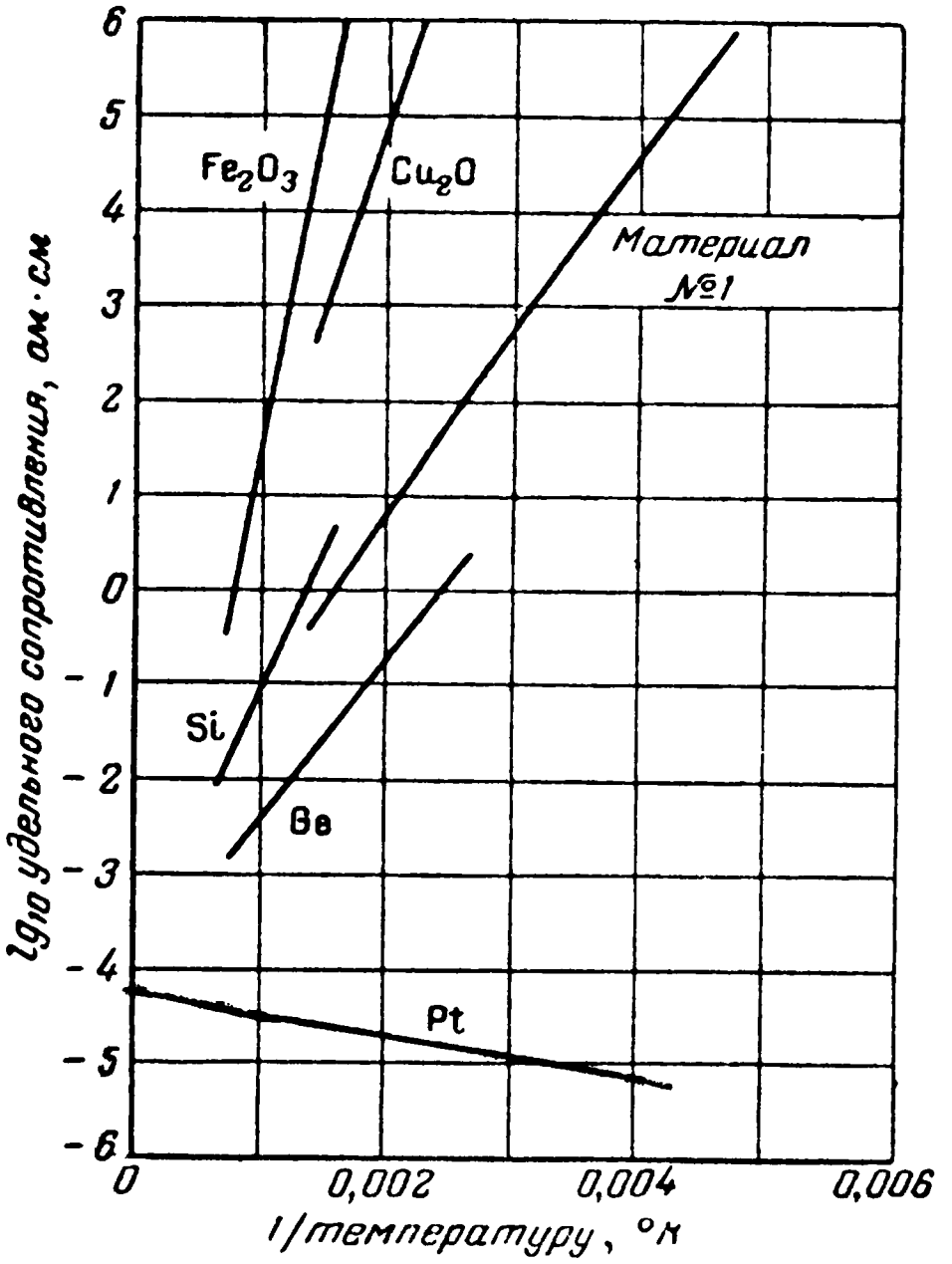


Рис. 2-4. Зависимости, показанные на рис. 2-3, получаются в виде прямых линий, если их представить как функции обратных температур.

гарифм удельного сопротивления, а по оси абсцисс — величину, обратную абсолютной температуре. Результаты приведены на рис. 2-4. Соответствующие аналитические выражения этих функций имеют вид:

$$\rho_T = \rho_\infty e^{B/T}, \qquad (2\text{-}1)$$

где ρ_T — удельное сопротивление при абсолютной температуре T;

ρ_∞ — удельное сопротивление при бесконечной температуре, полученное путем экстраполяции к точке, где $1/T = 0$; T выражено в абсолютных градусах;

B — коэффициент, численно равный тангенсу угла наклона данной прямой, умноженному на 2,303.

Теперь каждый полупроводник можно охарактеризовать величинами ρ_∞ и B. В уравнении (2-1) B играет ту же роль, что и работа выхода Ф (если отвлечься от несущественных численных коэффициентов) в формуле Ричардсона для термоэлектронной эмиссии. Действительно, как B, так и Ф определяют энергию активации.

Если покажется затруднительным оперировать бесконечной температурой в уравнении (2-1), то можно вместо него использовать уравнение

$$\rho_T = \rho_{T_0} e^{B(1/T - 1/T_0)}, \qquad (2\text{-}2)$$

где T_0 представляет некоторую относительную температуру, например 0° С (273° К).

Температурный коэффициент сопротивления α определяется как прирост величины удельного сопротивления, отнесенный к удельному сопротивлению, деленный на прирост температуры,

т. е.

$$\alpha = \frac{1}{\rho} \cdot \frac{d\rho}{dT}. \qquad (2\text{-}3)$$

Можно показать, что между α, B и T существует простая связь. Из уравнений (2-1) и (2-3) следует:

$$\alpha = -\frac{B}{T^2}. \qquad (2\text{-}4)$$

Хотя графики на рис. 2-4 представляются примерно прямыми линиями, детальное изучение показывает небольшое повышение наклона кривых по мере повышения температуры. Для разных материалов это повышение различно. Эта особенность имеет интересное теоретическое обоснование, но оно несущественно для инженерных целей.

2-2-1. Влияние примесей на удельные сопротивление и проводимость

Если построить график удельного сопротивления или удельной проводимости полупроводника в более широких пределах температур, чем использованные на рис. 2-4, то обнаружатся значительные отклонения от ли-

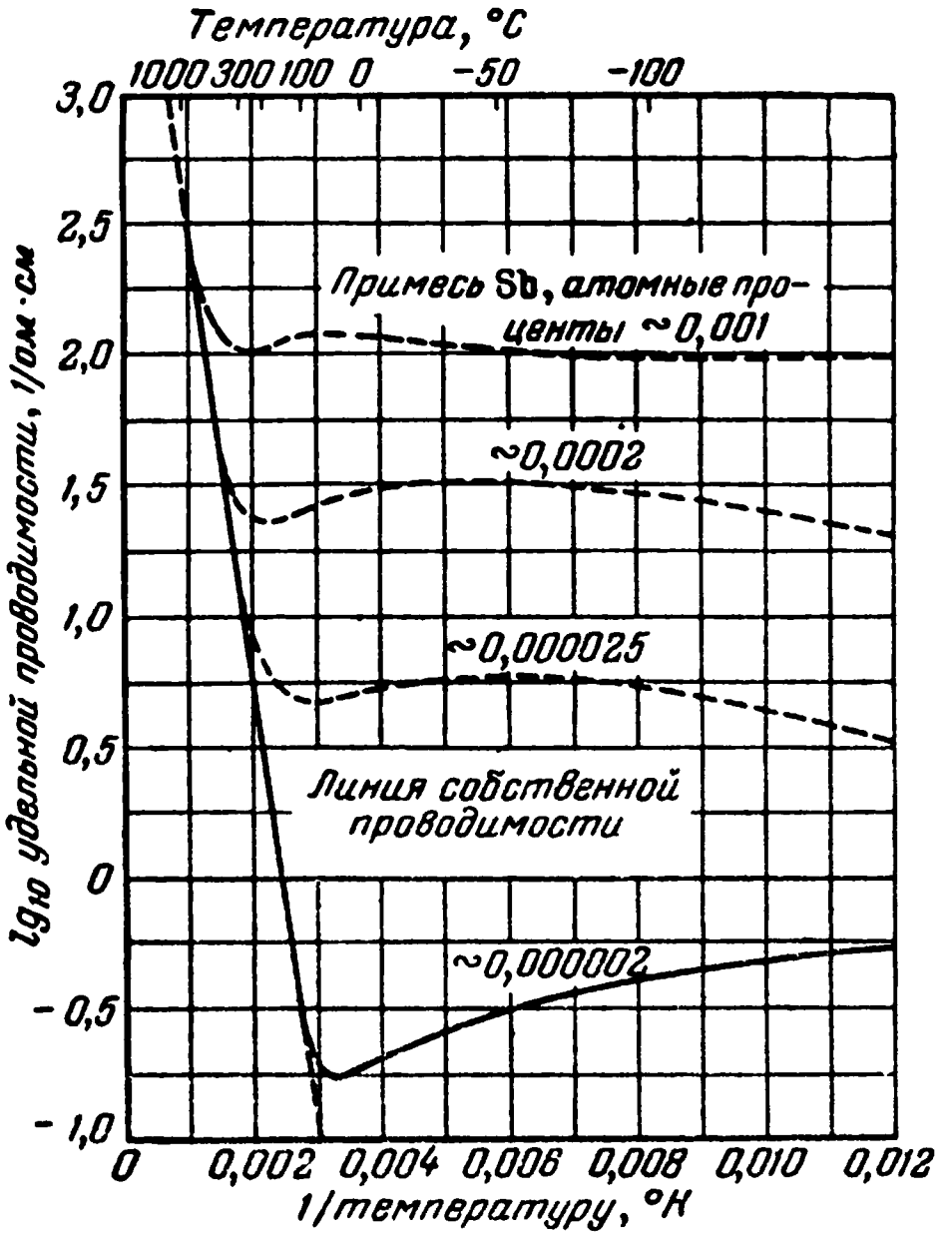


Рис. 2-5. При низких температурах удельная проводимость германия сильно зависит от количества примешанной сурьмы. При достаточно высокой температуре характеристики удельной проводимости всех образцов германия приближаются к общей характеристике, названной на графике линией собственной проводимости

нейной зависимости, выражаемой уравнением (2-1). На рис. 2-5 сплошной линией [Л. 1] показан график логарифма удельной проводимости (логарифм проводимости равен минус логарифму удельного сопротивления) германия для диапазона температур, более чем в 4 раза превышающего диапазон температур, указанный на рис. 2-4 для германия.

Эти данные получены для образца, содержавшего примесь примерно 0,000002% сурьмы, которая сообщает германию проводимость *n*-типа. Видно, что линейная зависимость имеет место только при температурах выше примерно 100° С. Исследование образца при любой температуре выше указанной показывает, что он имеет собственную проводимость. При более низких температурах полупроводник проявляет проводимость *n*-типа за счет примеси сурьмы.

Пунктирные кривые на этом графике показывают влияние примеси бóльших количеств сурьмы к тому же исходному материалу. Когда содержание сурьмы увеличивается, проводимость при низких температурах возрастает примерно пропорционально концентрации примеси. Подобным же образом при повышении содержания сурьмы повышается температура, при которой проводимость *n*-типа переходит в собственную проводимость. При достаточно высоких температурах все образцы ведут себя одинаково. Создается впечатление, что при высоких температурах возникает какой-то процесс, который парализует влияние примеси, сказывающееся при низких температурах. Это предположение правильно, как мы увидим позднее.

Можно также предположить, что, идя в другом направлении и удаляя последние следы сурьмы из исходного материала, можно заставить образец проявлять собственную проводимость при комнатной и даже при еще более низких температурах. Переход от проводимости, обусловленной примесью, к собственной проводимости, очевидно, зависит как от температуры, так и от количества примеси.

Некоторые примеси эффективнее увеличивают проводимость полупроводника, чем другие. В случае только что рассмотренной примеси сурьмы к германию считается, что каждый атом сурьмы добавляет один свободный электрон к числу носителей электричества, находящихся внутри германия. Примесь одной десятитысячной бора к кремнию может повысить его проводимость при комнатной температуре более чем в миллион раз. С другой стороны, некоторые примеси могут присутствовать в концентрациях до 1%, не влияя заметным образом на проводимость. Далее, эффективность примеси зависит от степени ее дисперсности в массе полупроводника.

Если примесь присутствует в виде истинного твердого раствора, когда отдельные атомы или ионы примеси распределены по кристаллической решетке полупроводника, эффективность примеси наибольшая. Если же примесь содержится в виде больших агрегатов, комков или включений, ее влияние значительно ослабляется.

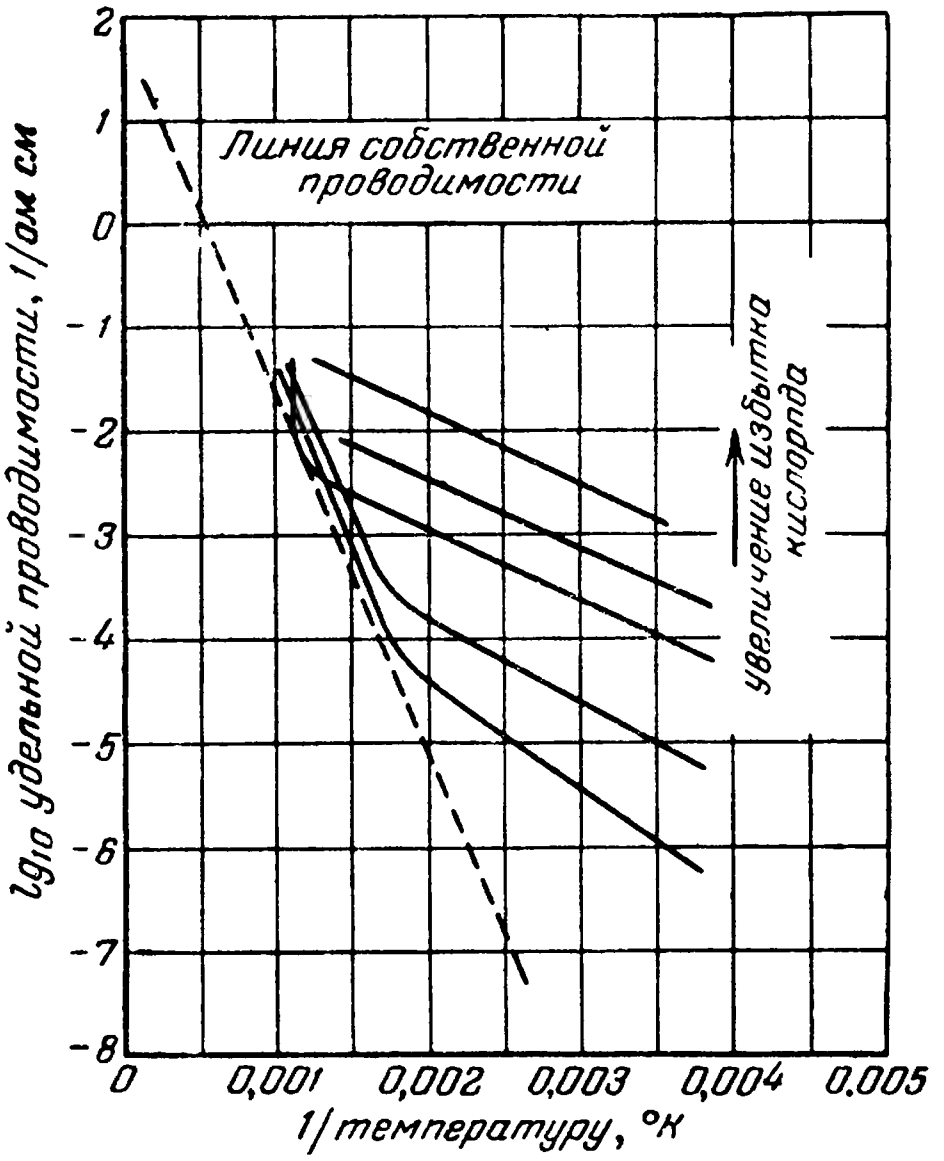


Рис. 2-6. Характеристики удельной проводимости закиси меди при разных содержаниях избытка кислорода.

Поскольку термическая обработка играет значительную роль в распределении примеси внутри вещества, нередко можно управлять свойствами полупроводника надлежащей тепловой обработкой — отжигом, сплавлением, закалкой.

Очевидно, что приблизительно прямые линии на рис. 2-4 являются лишь соответствующими собственной проводимости участками более сложных кривых, показанных на рис. 2-5.

Проводимость полупроводника может зависеть и от других форм «примесей», помимо разобранного случая введения атомов инородного элемента в виде твердого раствора. В частности, в оксидных полупроводни-

ках «примесью» может являться избыток или недостаток кислорода относительно точных пропорций, необходимых для образования данных видов окислов. Так, в закиси меди всегда находится небольшой избыток кислорода и величина избытка в известной степени может регулироваться нагревом образца в среде с соответствующим парциальным давлением кислорода. На рис. 2-6 дано семейство кривых логарифма удельной проводимости как функций $1/T$ для серии образцов закиси меди, имеющих различное количество избыточного кислорода [Л. 2]. Здесь мы видим особенности, уже описанные для серии смесей сурьмы с германием: при низких температурах — зависимость проводимости от содержания примеси, при высоких — переход к собственной проводимости, причем ее величина примерно одинакова для всех образцов.

Строго рассуждая, когда мы говорим «избыток кислорода», мы подразумеваем недостаток меди. Атом или ион кислорода слишком велик, чтобы поместиться между узлами кристаллической решетки закиси меди. Вместо этого некоторые ионы меди отсутствуют на их нормальных местах в решетке. Оставшиеся пустые места при этом действуют как «примеси», вызывающие проводимость p-типа.

2-3. ВЫПРЯМИТЕЛЬНЫЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ

Если для обычного сопротивления построить график падения напряжения как функцию проходящего через него тока, то получится прямая линия, проходящая через начало координат, наклон которой численно равен сопротивлению. Так ведет себя бóльшая часть встречающихся сопротивлений. Их характеристики линейны относительно тока и напряжения; они подчиняются закону Ома и поэтому называются омическими (активными) сопротивлениями. Однако имеются проводники, электрические характеристики которых нелинейны. Более или менее знакомыми примерами таких проводников являются электрическая дуга (имеющая отрицательное сопротивление в некоторой части характеристики) и вакуумный диод, работающий в режиме пространственного заряда (ток равен нулю при отрицательном потенциале анода, изменяется пропорционально потенциалу анода в степени $^3/_2$ для малых положительных потенциалов

анода и остается примерно постоянным в режиме насыщения). Нелинейные сопротивления играют в электронике не меньшую роль, чем омические.

Интересным и полезным свойством полупроводников является то, что при контакте полупроводника с надлежаще выбранным металлом или другим полупроводником можно получить нелинейное сопротивление. Полупроводниковые приборы, работа которых основана на нелинейности характеристик таких контактов, называются варисторами или переменными сопротивлениями.

Различные виды варисторов можно подразделить на два основных класса: симметричные и несимметричные варисторы. Симметричные варисторы имеют нелинейные характеристики, одинаковые для любой полярности приложенного напряжения, тогда как характеристики несимметричных варисторов, помимо нелинейности, зависят от полярности приложенного напряжения. В последний класс входят диоды и контактные выпрямители всех типов.

Несимметричные варисторы обычно далее подразделяются на плоскостные и точечные выпрямители. Плоскостные выпрямители используются там, где приходится работать со сравнительно большими токами, так как величина тока, пропускаемого выпрямителем, зависит от площади контакта. Плоскостные выпрямители находят основное применение как выпрямители для зарядки аккумуляторов, для дуговых сварочных аппаратов, для выпрямителей питания цепей накала и анодов ламп в электронных устройствах. Точечные выпрямители, представляющие собой детекторы с контактом в виде тонкого острия, применяются там, где нужно получить нелинейную электрическую характеристику, а величина пропускаемого тока не играет роли. Эти приборы используются для модуляции, демодуляции [1] и детектирования при слабых токах В дальнейших главах мы рассмотрим такие приборы, их свойства, теорию и физические основы их работы.

---

[1] Термины «детектирование» и «демодуляция» в принципе равнозначны, но несколько различаются по области их применения Согласно установившейся отечественной терминологии под детектированием обычно понимается процесс выделения огибающей из промодулированного высокочастотного сигнала в приемных устройствах радиосвязи с помощью простейшей диодной схемы; аналогичный процесс в технике проводной связи называется демодуляцией, при-

2-4. ФОТОЭЛЕКТРИЧЕСКИЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ

Устройства, в которых используются изменения свойств светочувствительного элемента при его освещении, называются фотоэлектрическими элементами, или просто фотоэлементами. Согласно точному смыслу этого определения зеленые листья и фотографические пластинки тоже надо называть фотоэлементами, поскольку они действуют благодаря фотохимическим изменениям при воздействии света. Чтобы привести это определение в соответствие с тем, что мы подразумеваем под словом «фотоэлемент», нужно ввести в определение несколько уточнений. Нужно оговорить, что изменения должны касаться электрических свойств элемента и что они должны возникать в результате фотофизических, а не фотохимических процессов.

Существуют три разных фотофизических эффекта. Первый из них заключается в испускании электронов с поверхности фотокатода в окружающее пространство. Этот процесс называется фотоэлектронной эмиссией, и бóльшая часть встречающихся на практике фотоэлементов работает на этом принципе. Электроны, испускаемые фотокатодом в количестве, пропорциональном падающему световому потоку, улавливаются положительно заряженным анодом, расположенным поблизости, и создают электрический ток во внешней цепи прибора.

Во втором случае электрическая проводимость светочувствительного полупроводникового элемента увеличивается при его освещении. Это явление называется внутренним фотоэффектом, а устройства, в которых оно используется, называются фотосопротивлениями. Повышение проводимости, вызываемое светом, приводит к увеличению тока, если в цепи такого фотоэлемента включена батарея.

Наконец, при фотогальваническом эффекте свет, падающий на выпрямительный контакт или спай, вызывает

чем по большей части имеется в виду метод передачи сигнала на одной боковой полосе с подавлением несущей. Используемые в этом случае устройства, как правило, собраны по более или менее сложной схеме (балансной, кольцевой и т. д) на нескольких вентилях и для работы требуют подачи опорной частоты от постороннего генератора. При этом выбор того или иного термина определяется не принципом действия устройства, а областью его применения. *Прим. перев.*

появление э. д. с. Светочувствительный элемент в электрическом экспонометре для фотографии представляет собой фотоэлемент этого типа, так же как и солнечная батарея. Эти устройства не нуждаются во внешних источниках тока, действуя сами при облучении их светом как батареи.

Все выпускаемые промышленностью полупроводниковые фотоэлементы представляют собой либо фотосопротивления, либо фотогальванические элементы. Явление фотоэлектронной эмиссии в полупроводниках еще не нашло применения в коммерчески используемых устройствах, однако оно используется как экспериментальный метод для изучения свойств поверхностей полупроводников и распределения электронов в полупроводниках по энергиям. Как мы увидим в гл. 22, между спектральной чувствительностью электровакуумного фотоэлемента и работой выхода материала его катода имеется определенная зависимость. Для большинства полупроводников работа выхода слишком высока и соответствующий участок длин волн оказывается трудно используемым. В качестве фотоэлектронных катодов поэтому применяются почти исключительно металлы с низкой работой выхода [1].

Полупроводниковые фотосопротивления и фотогальванические элементы оказываются чувствительными к более широко используемому участку оптического спектра. В то время как спектральная чувствительность многих электровакуумных фотоэлементов ограничена зеленым, синим, фиолетовым и ультрафиолетовым участками спектра, полупроводниковые фотосопротивления и фотогальванические элементы чувствительны главным образом к красному и инфракрасному участкам спектра. Если учесть, что бо́льшая часть фотоэлементов работает с искусственными источниками света, характеристика излучения которых приблизительно совпадает с характеристикой абсолютно черного тела, имеющей максимум в области длин волн порядка микрона, то преимущества полупроводниковых приборов в этой области выступают весьма четко. Наиболее привлекательная особенность

---

[1] Автор имеет в виду, очевидно, такие полупроводники, как германий и кремний В современных электровакуумных фотоэлементах применяются полупроводники в виде соединений, например, сурьмы и цезия или висмута и цезия. *Прим. ред.*

полупроводниковых фотоэлементов заключается в том, что некоторые типы фотосопротивлений имеют чувствительность, в десятки и сотни раз бóльшую, чем лучшие электровакуумные фотоэлементы. Недостатками полупроводниковых фотоэлементов являются нестабильность их параметров и небольшой срок службы, в особенности при несколько повышенных температурах. Очень часто они оказываются также слишком инерционными, так как не в состоянии воспринимать быстрые изменения светового потока.

2-5. ДВУПОЛЯРНАЯ ПРОВОДИМОСТЬ В ПОЛУПРОВОДНИКАХ

Внутренность полупроводника представляет собой совокупность атомов, электронов, положительных дырок и пустых мест. Частицами, определяющими электрические свойства полупроводника, являются содержащиеся в нем электроны и положительные дырки. Эти частицы своим наличием и движением в кристаллической решетке под влиянием тепловых, электрических, оптических и магнитных воздействий обусловливают проводимость, выпрямительное действие, фотоэффект и другие свойства полупроводника.

В каждом полупроводнике содержится некоторое количество обоих указанных видов частиц. Как было выяснено выше, в полупроводниках *n*-типа численно преобладают электроны, в полупроводниках *p*-типа преобладают дырки. В полупроводниках с собственной проводимостью электронов и дырок примерно поровну. Чтобы дать некоторое представление о количестве частиц, можно заметить, что образец германия с проводимостью *n*-типа, содержащий умеренное количество примеси, сбусловливающей указанный тип проводимости, при комнатной температуре имеет около 10^{16} свободных электронов в 1 $см^3$ и около 10^{10} дырок в 1 $см^3$. При повышении температуры концентрация частиц обоих типов повышается. Относительный прирост содержания всегда выше для тех частиц, которые находятся в меньшинстве. Так, если температура указанного выше образца германия повысится настолько, что концентрация электронов в нем удвоится и дойдет до $2 \cdot 10^{16}$ электронов в 1 $см^3$, то концентрация дырок повысится до 10^{14} дырок в 1 $см^3$, т. е. в 10 000 раз. По мере повышения температуры кон-

центрация частиц, находящихся в меньшем количестве (неосновных носителей), достигает концентрации частиц, находящихся в большинстве (основных носителей), и образец приобретает собственную проводимость; для более высоких температур концентрации обоих видов частиц увеличиваются примерно в одинаковой степени.

Поскольку удельная проводимость полупроводника зависит, помимо прочих факторов, от концентрации подвижных частиц, способных участвовать в создании тока, из предыдущего примера легко понять, почему удельная проводимость полупроводника повышается при увеличении температуры.

В то время как нормальная концентрация неосновных носителей в германии составляет при комнатной температуре от 10^9 до 10^{10} частиц в 1 $см^3$, можно значительно повысить эту величину, эмитируя частицы в полупроводник из соответствующего контакта, прижатого к его поверхности и имеющего подходящий потенциал. Эти частицы являются посторонними в том смысле, что они временно находятся в местах, где их обычно не бывает, подобно электронам в вакууме. Подобно последним, их можно перемещать внутри полупроводника с помощью надлежащим образом направленных полей, вводить или отсасывать с помощью электродов соответствующей формы, находящихся под соответствующим потенциалом, и вообще можно заставить их вести себя внутри полупроводника во многом подобно тому, как они ведут себя в вакууме.

Способность полупроводника одновременно содержать подвижные заряженные частицы противоположных знаков определяет его свойство двуполярной проводимости. На этом свойстве основано действие транзисторов и подобных им приборов.

Посторонние частицы, введенные в полупроводник, живут в нем не очень долго. Введение их нарушает существовавшее ранее состояние равновесия между концентрациями различных видов частиц, и полупроводник стремится восстановить равновесие, вызывая исчезновение лишних частиц. Это происходит путем рекомбинации дырок и электронов. Когда дырка и электрон встречаются при перемещении внутри кристаллической решетки полупроводника, они взаимно аннигилируются и исчезают из числа создающих ток носителей заряда.

Чем выше избыток концентрации неосновных носителей, тем выше вероятность такой встречи и тем выше скорость восстановления. Типичная длительность существования неосновных носителей в германии — от нескольких единиц до нескольких сотен микросекунд. Таким образом, если инжектированные частицы должны быть использованы для какой-либо полезной цели, то это должно быть сделано быстро: если нужно переместить частицы, то расстояние должно быть мало, иначе частицы будут уничтожены раньше, чем они попадут в место назначения. Следовательно, рабочая часть транзистора должна иметь малые геометрические размеры, обычно несколько тысячных сантиметра. Роль длительности существования неосновных носителей в действии транзисторов рассмотрена в гл. 15—25 этой книги.

2-6. ТЕРМОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

Основные сведения о механизме проводимости твердых тел можно получить, изучая термоэлектрические явления. В 1821 г. Зеебек нашел, что если присоединить к двум концам металла куски другого металла и нагреть один из контактов, то в разрыве цепи второго

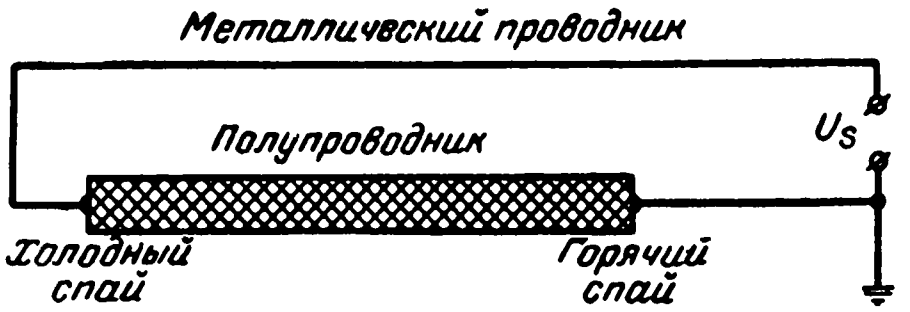


Рис. 2-7. Цепь для измерения термо- э. д. с. полупроводника.

металла появится электрическое напряжение. На рис. 2-7 показана схема цепи для измерения термо- э. д. с. полупроводника. Если присоединить металлические контакты к двум концам полупроводникового стержня или бруска и одно из соединений будет находиться при более высокой температуре, чем другое, то между контактами появится разность потенциалов. Эта термо- э. д. с. возникает частично за счет того, что внутри полупроводника за каждую секунду от горячего к холодному концу диффундирует больше основных носителей заряда, чем в противоположном направлении, образуя таким обра-

зом разность потенциалов между концами. Это напряжение возрастает до тех пор, пока вызванный им обратный ток не уравновесит ток диффузии и не будет достигнуто динамическое равновесие. Другая часть этой термо-э. д. с. создается за счет контактной разности потенциалов между полупроводником и металлом на двух контактах различного температурного коэффициента.

Как следует из схемы, представленной на рис. 2-7, холодный спай заряжается отрицательно, если проводимость обусловливается в основном электронами; если проводимость создается дырками, то холодный спай становится положительным. Если величина напряжения U_s не очень велика, она пропорциональна разности температур между горячим и холодным спаями. Исследуя полярность термо- э. д. с., можно сделать вывод о типе проводимости образца. Далее, по величине термо- э. д. с. можно определить концентрацию носителей заряда в образце. Теоретическое обоснование этих выводов дано в гл. 23.

2-7. ЭФФЕКТ ХОЛЛА В ПОЛУПРОВОДНИКАХ

Имеется еще один путь для получения основных сведений о видах и концентрациях носителей заряда — исследование эффекта Холла. В 1879 г. Е. Холл [Л. 3] обнаружил, что если вдоль металлической полоски пропустить ток и поместить ее в магнитное поле, направленное перпендикулярно току, то в направлении поперек полоски возникнет разность потенциалов. Рассмотрим прямоугольную пластинку, расположенную в плоскости $Z=0$, с ребрами, параллельными осям X и Y, как показано на рис. 2-8. Если в направлении оси X протекает ток I_x от внешнего источника и если пластинка помещена в магнитное поле H_y, направленное по оси Y, то между зажимами A и B возникает разность потенциалов U_z. Величина ее в вольтах определяется уравнением

$$U_z = \frac{10^{-8} R_H I_x H_y}{d}, \qquad (2\text{-}5)$$

где I_x — ток, *а*; H_y — магнитное поле, *э*; d — ширина пластинки, *см* и R_H — коэффициент пропорциональности, называемый постоянной Холла, в $см^3/к$. Величина R_H зависит от материала пластинки, содержания примесей и температуры.

Эффект Холла в полупроводниках *n*-типа возникает вследствие того, что магнитное поле искривляет путь электронов в пластинке при их движении от одного конца к другому. Таким образом, вдоль того края проводника, в сторону которого изгибаются траектории электронов, образуется избыточный отрицательный заряд. Этот заряд растет до тех пор, пока образуемое им электрическое поле своим действием не парализует искривляющее влияние магнитного поля. Как можно видеть из

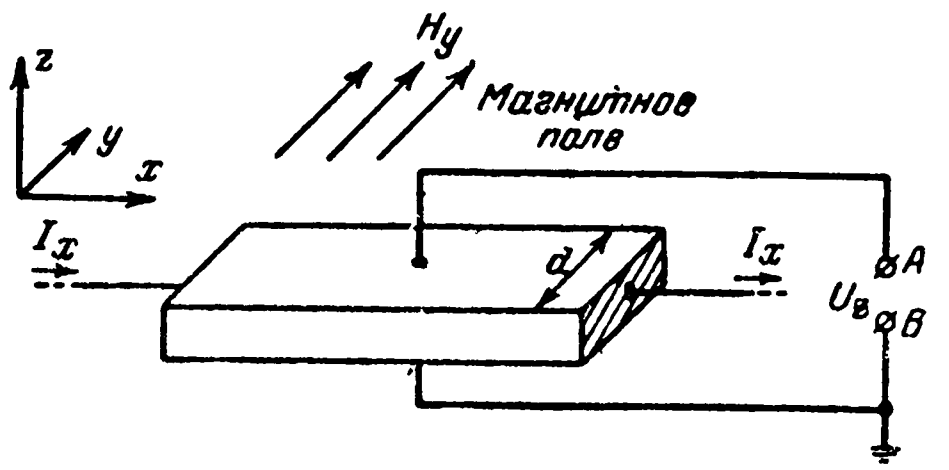


Рис. 2-8 При прохождении тока в поперечном магнитном поле возникает э. д с. Холла, направленная под прямым углом к магнитному полю и току

уравнения (2-5), величина поперечного напряжения пропорциональна как величине тока, так и напряженности магнитного поля. При эффекте Холла электроны ведут себя подобно воде в реке, в русле которой установлены перегородки, не доходящие до поверхности и расположенные так, чтобы отклонять воду к одному берегу. Уровень воды возле этого берега выше, чем у другого, ровно настолько, чтобы поперечный поток за счет этой разности уровней в точности компенсировал отклоняющее действие перегородок.

При условиях, показанных на рис. 2-8, зажим *A* должен быть отрицательным относительно зажима *B*. Так будет в том случае, когда ток в металле или полупроводнике создается потоком свободных электронов. Однако для многих веществ в подобном опыте зажим *A* оказывается заряженным положительно. Это на первый взгляд аномальное явление вызывало большое замешательство в научных кругах, пока не возникло представление о положительных дырках. В настоящее время, однако, это явление объясняется очень просто: для проводников *p*-типа, в которых ток создается потоком поло-

жительных дырок, поперечное напряжение за счет эффекта Холла должно иметь противоположный знак по сравнению со случаем переноса заряда электронами. Таким образом, можно определить тип проводимости полупроводника, определив полярность э. д. с. получающейся в опыте с эффектом Холла.

Кроме того, сочетая данные, полученные с помощью эффекта Холла, с данными, полученными из измерения удельной проводимости, можно получить сведения о концентрации и подвижности носителей заряда любого конкретного вещества или образца. Наблюдая зависимость постоянной Холла и величины удельной проводимости от температуры и содержания примесей в образце, можно уточнить наши сведения о природе и механизме проводимости. Поскольку эти выводы являются в значительной степени теоретическими и понимание их зависит от глубокого знания внутренних процессов, происходящих в твердых телах, дальнейшее рассмотрение эффекта Холла продолжено лишь в гл. 24.

ЗАДАЧИ

1. Покажите, что удельная проводимость имеет физическую размерность *мо/см*.

2 На основании данных, приведенных на рис. 2-3, определите температурный коэффициент сопротивления для платины при 0° С.

3. Сравните найденную в п. 2 величину с температурным коэффициентом сопротивления для термисторного материала № 1 при той же температуре, определив его любым путем из данных, приведенных на рис. 2-3 или 2-4. (Обратите внимание, что на рисунках использованы десятичные логарифмы, а в выводах уравнений, приведенных в тексте, — натуральные.)

4. Рассчитайте сопротивление при комнатной температуре (300° К) элемента, изготовленного из термисторного материала № 1, имеющего квадратное сечение со стороной 1 *мм* и длину 3 *см*. Какова проводимость этого элемента?

5. Элемент из задачи 4 помещен в магнитное поле с напряженностью 1 000 *э* и через него пропускается ток 10 *ма*. Какова будет величина э. д. с. Холла между двумя точками, расположенными точно друг против друга на противоположных сторонах элемента? Постоянная Холла для термисторного материала № 1 равна 40 $см^3/к$. Элемент ориентирован так, что направления магнитного поля, тока и измерения э. д. с. Холла взаимно перпендикулярны.

6. Можете ли Вы объяснить, почему для полупроводника *n*-типа направление э. д. с. Холла на рис. 2-8 должно быть таким, что точка *A* отрицательна, а точка *B* — положительна, тогда как для полупроводников *p*-типа точка *A* будет положительной, а *B* — отрицательной?

7. Каков должен быть прирост температуры, чтобы сопротив-

ление элемента из закиси меди уменьшилось в 2 раза, начиная с 0° С? Начиная с 200° С?

8 Электрон из атома донорной примеси может рекомбинировать с дыркой из атома акцепторной примеси и нейтрализовать ее. Как будет меняться удельная проводимость, если взять полупроводник n-типа, имеющий фиксированное число доноров в 1 $см^3$, и вводить все больше и больше акцепторов в кристалл? Начертите качественный график удельной проводимости как функции концентрации добавленных акцепторов.

9. Можете ли Вы дать физическое толкование того факта, что все кривые на рис. 2-4 стремятся к общей точке пересечения ρ_∞ при приблизительно 10^{-4} $ом \cdot см$? Можно ли предположить, что возможно создание материала с нулевым температурным коэффициентом сопротивления? Какой можно ожидать величину удельного сопротивления для такого материала? Если был бы открыт металл, имеющий удельное сопротивление при комнатной температуре равное 10^{-8} $ом \cdot см$, то какова была бы величина его температурного коэффициента сопротивления при комнатной температуре?

ЛИТЕРАТУРА

1. Pearson G L, упомянуто в работе K. Lark-Horovitz, Conductivity in Semiconductors, Electrical Engineering, 1949, vol. 68, p 1047.

2 Jusé und Kurtschatow, Zur elektrischen Leitfähigkeit von Kupferoxydul, Physikalische Zeitschrift der Sowjetunion, 1932, Bd. 2, S. 453.

3. Hall E. H, On a New Action Of the Magnet on Electric Currents, American Journal of Mathematics, 1879, vol. 2, p. 287.

Полупроводники в науке и технике, под ред. А. Ф. Иоффе, Изд. АН СССР, 1957.

Иоффе А. Ф, Физика полупроводников, Изд. АН СССР, 1957.

Гуров В. С., Полупроводники в технике и быту, Изд. «Московский рабочий», 1958

Иоффе А. Ф., Полупроводники и их применение, Изд. АН СССР, 1956.

ГЛАВА ТРЕТЬЯ

ТЕРМОСОПРОТИВЛЕНИЯ (ТЕРМИСТОРЫ) [Л. 1]

3-1. ОБЩИЕ СВЕДЕНИЯ

Полупроводниковые термосопротивления (термисторы) представляют собой приборы, использующие высокий температурный коэффициент сопротивления полупроводников. Способность изменять сопротивление, малые размеры и механическая прочность этих приборов

обусловили их широкое внедрение в аппаратуру связи уже в первые 5 лет после их появления. Термисторы могут быть изготовлены самых различных размеров и форм и с таким разнообразием термических и электрических характеристик, что их можно применять в любых устройствах подобного рода. Это многообразие форм является еще одной из причин широкого применения термисторов.

Первоначально термисторы использовались в термометрии и для компенсации температурных изменений, влияющих на работу различных устройств. Поскольку термистор имеет большой отрицательный температурный коэффициент сопротивления, включение термистора в цепь из металлических проводников, имеющую положительный температурный коэффициент, может при надлежащем подборе характеристик термистора сделать характеристики цепи почти не зависящими от температуры в диапазоне до 100° С. В дальнейшем термисторы стали применять в линейных усилителях проводных линий дальней связи для обеспечения плавной автоматической регулировки усиления с целью компенсации температурных изменений характеристик присоединенного участка линии. Такие скомпенсированные линии связи имеют почти постоянные характеристики независимо от суточных и сезонных изменений местной температуры. Во время и после войны термисторы использовались в измерителях мощности СВЧ. Высокий температурный коэффициент обеспечивает термисторам в этом случае применения чувствительность на один порядок большую, чем у термопар. Джоулево тепло, выделяемое в термисторе при прохождении через него тока, приводит к нелинейности его характеристики. Эта нелинейность используется в ограничителях сигналов, стабилизаторах напряжения и усилителях на отрицательных сопротивлениях. Поскольку термисторный элемент обладает тепловой инерцией, характеристики изменения его тока во времени могут быть использованы в схемах выдержки времени.

Помимо указанных выше применений термисторов в технике связи, они используются в метеорологии, медицине, науке и промышленности. Термисторы применяются как датчики температуры в радиозондах. В медицине они используются для внутривенной термометрии. Повсеместно употребляются термисторные вакуумные

манометры, и в промышленности имеется множество термисторных расходомеров. Термисторные элементы в форме тонких листков, имеющих малую тепловую инерцию, используются в спектроскопии как индикаторы излучения. В военном деле термисторы применяются для обнаружения ночью на значительном расстоянии нагретых объектов, как, например, дымовые трубы кораблей, моторы, и даже люди. Термисторы входят в конструкцию самолетных вариометров (индикаторов скорости подъема и снижения).

Тепловые свойства термисторов известны уже в течение более 100 лет, но тем не менее термисторы не изготовлялись и не использовались до 30-х годов. Ранние опыты с полупроводниками принесли разочарование из-за плохой воспроизводимости результатов и электрической нестабильности изготовленных образцов[1]. Один экспериментатор говорил: «Удельное сопротивление германия равно одной тысячной омо-сантиметра». Его собеседник возражал: «Нет, оно равно 20 омо-сантиметрам; я только что измерил его». Лишь после создания современной теории твердых тел и разработки соответствующей технологии была по-настоящему понята огромная зависимость свойств полупроводников от природы и распределения примесей и получена возможность контролировать эти факторы настолько, чтобы изготовлять воспроизводимые стабильные полупроводниковые приборы.

3-2. ИЗГОТОВЛЕНИЕ ТЕРМИСТОРОВ ПРЯМОГО ПОДОГРЕВА

Термисторы изготовляются в виде бусинок, стержней, дисков, шайб и пленок. Каждый термистор может быть выполнен различных размеров и из разнообразных полупроводниковых материалов для получения желательных электрических характеристик. На рис. 3-1 показаны фотографии нескольких типов термисторов, демонстрирующие разнообразие форм и размеров. На рис. 3-2 приведены эскизы конструктивного выполнения этих прибо-

---

[1] Конечно, дело не только в нестабильности и плохой воспроизводимости ранних полупроводниковых приборов, но и в неподготовленности к их использованию всей техники того времени и в том, что ее развитие неизбежно должно было базироваться на приборах, физические процессы в которых были лучше известны. *Прим. ред.*

ров и показано размещение термисторных элементов внутри оболочки.

Термисторы могут быть изготовлены любым методом, позволяющим отформовать и снабдить металлическими контактами полупроводниковый элемент нужных размеров. Наиболее распространенный метод заключается

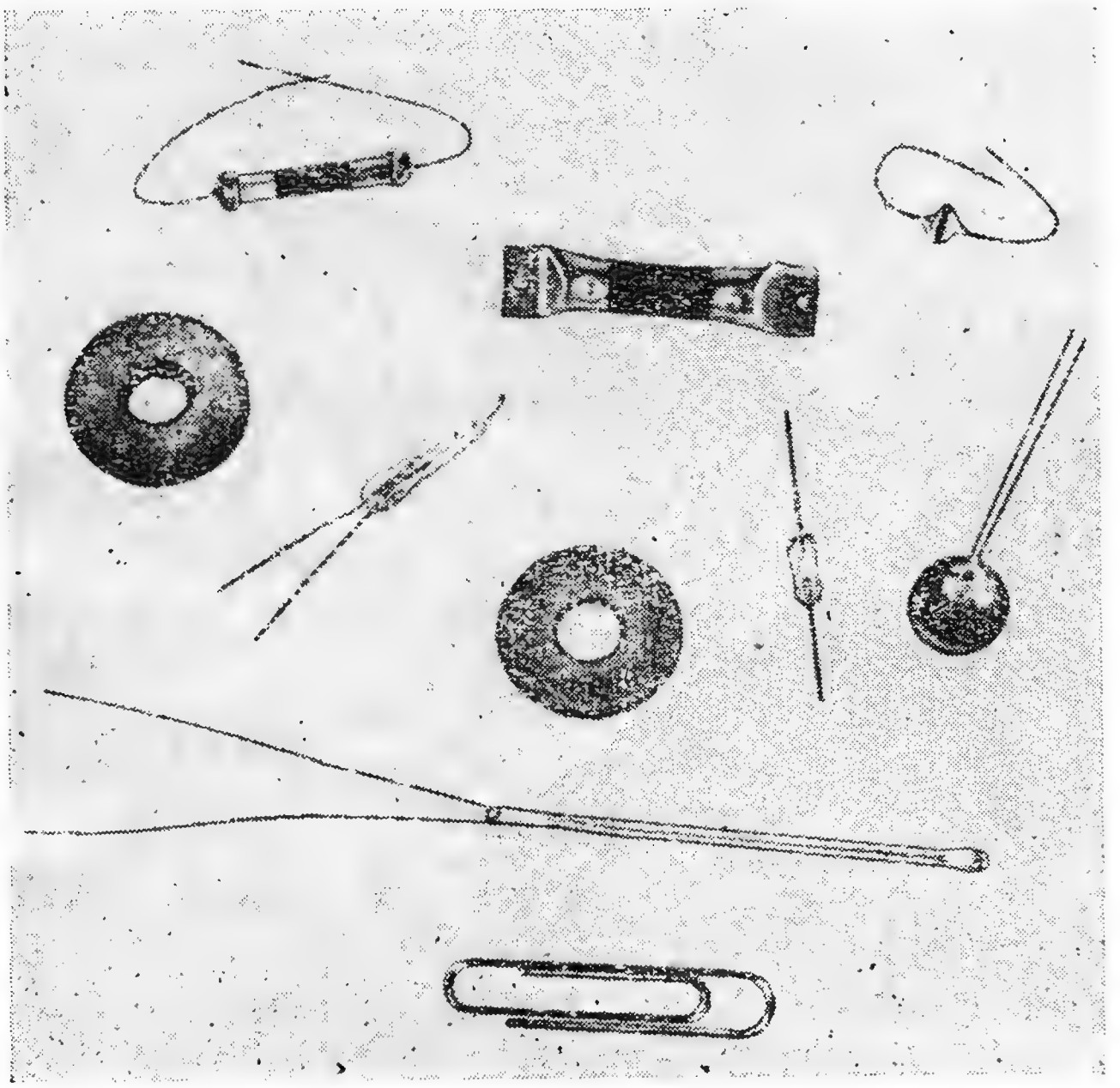
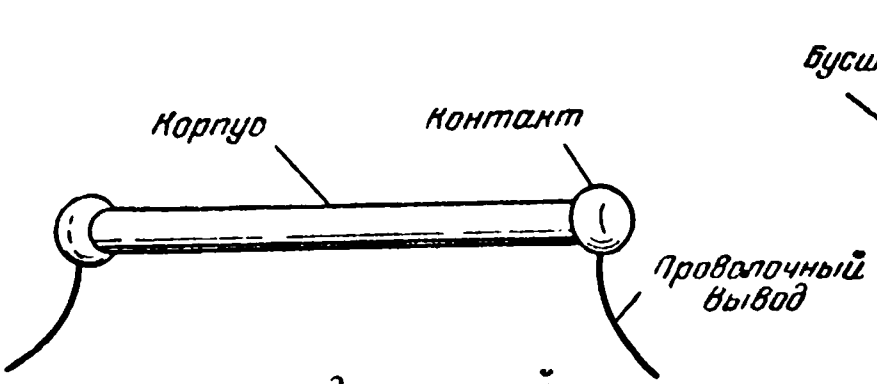


Рис 3-1. Термисторы разных размеров, формы и с различными защитными оболочками

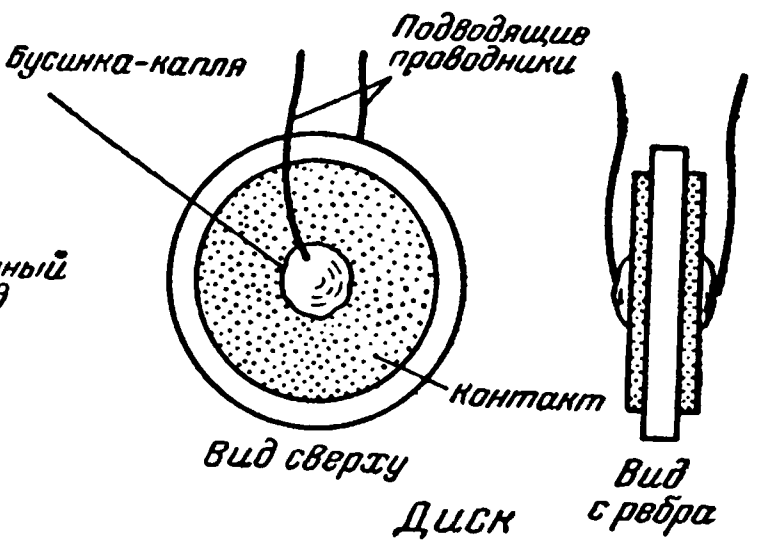
в приготовлении химическим способом желательного полупроводника или смеси полупроводников в порошкообразном виде, формовке из него элемента нужного размера и формы и нагреве его до достаточно высокой температуры, при которой зерна порошка сплавляются вместе в прочную, плотную, однородную массу. Этот процесс подобен используемому в порошковой металлургии. Металлические контакты обычно выполняются путем нанесения контактных поверхностей золотой, серебряной или платиновой пастой, подобной применяемой для росписи

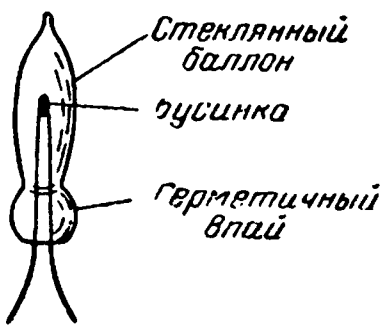
керамических изделий, и последующего обжига пасты для спекания в сплошную металлическую пленку, к которой могут быть припаяны проволочные выводы.

Бусинковые термисторы изготовляются натягиванием двух тонких проволочек из платинового сплава параллельно друг другу на расстоянии от пяти до десяти диа-

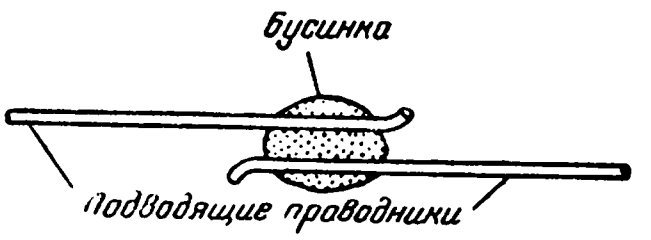


Цилиндрический стержневой термистор

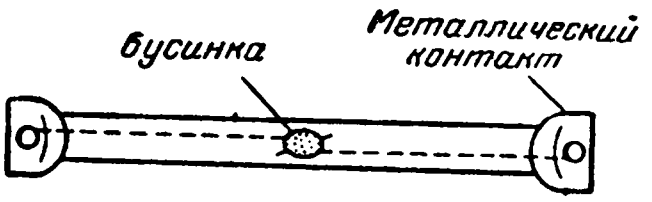




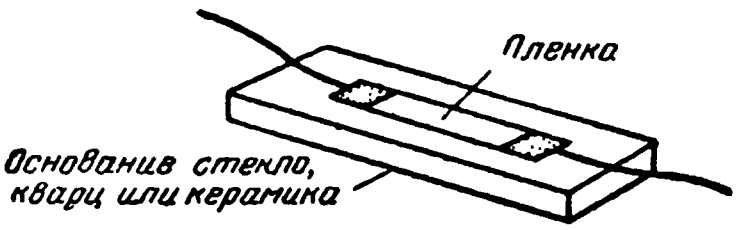
Бусинка в стеклянной запаянной трубке



Увеличенная деталь бусинки и проволочных держателей-выводов



Полый фибровый патрон для бусинкового термистора



Пленочный термистор

Рис. 3-2. Конструктивное выполнение некоторых термисторов, показанных на рис. 3-1.

метров. Затем на них наносится небольшая капля пасты из смеси порошка полупроводника с подходящим органическим биндером (связующим веществом). Поверхностное натяжение стягивает эту массу в почти сферическую бусинку. На проволочках располагается от 10 до 20 таких бусинок на одинаковых расстояниях. Далее бусинки просушиваются и слегка нагреваются для схватывания биндера и придания пасте достаточной прочности, чтобы можно было брать бусинку за проволочки.

После этого бусинки спекаются в печи. Частицы порошка стягиваются вокруг проволочек и создают плотный и постоянный контакт с ними. Затем проволочки обрезаются и образуют индивидуальные выводы. Диаметры бусинок могут быть от 0,15 до 1,5 *мм* при проволочках диаметром от 0,025 до 0,15 *мм*. В некоторых термисторах такого типа бусинка вплавлена в стеклянную оболочку. В других случаях она заключена в вакуумный или заполненный инертным газом стеклянный баллончик. Некоторые термисторы бусинкового типа заключены в цилиндрические фибровые кожухи с навальцованными на концах металлическими выводами.

Термисторы в виде стержней изготовляются смешиванием размельченного полупроводникового материала с органическим биндером и растворителем, продавливанием смеси сквозь глазок, просушкой, нарезкой на отрезки нужной длины, выжиганием биндера и, наконец, спеканием при высокой температуре. Контакты накладываются, как описано выше. Стержни делаются диаметром от 0,8 до 6,4 *мм* и длиной от 1,5 до 50 *мм*. Термисторы в виде дисков и шайб изготовляются подобным же образом прессовкой в формах порошка с биндером. Диаметр дисков от 1,5 до 50 *мм*, толщина от 0,8 до 6,4 *мм*.

Термисторы в форме пленок изготовляются смешиванием порошкообразного полупроводника с подходящим биндером до сметанообразной консистенции, нанесением этой массы тонким однородным слоем на гладкую стеклянную поверхность, просушиванием, снятием пленки и нарезанием ее на листочки желательной формы и размера, которые обжигают при температуре спекания на гладкой керамической поверхности, служащей основанием. Контакты накладываются, как описано выше. Возможные размеры листочков: толщина — 0,01—0,04 *мм*, длина — 1—10 *мм*, ширина — 0,2—1,0 *мм*.

Указанные выше геометрические размеры оказались наиболее удобными, но они ни в коей мере не исчерпывают возможностей современной технологии.

3-3. ФИЗИЧЕСКИЕ СВОЙСТВА САМОНАГРЕВАЮЩИХСЯ ТЕРМИСТОРОВ

Самонагревающийся термистор представляет собой двухполюсник, сопротивление которого определяется выбранной температурой. В свою очередь эта темпера-

тура зависит от температуры окружающей среды и джоулевого тепла, выделяемого внутри элемента за счет протекающего через него тока подогрева. При работе в переходных режимах температура элемента зависит также от предшествующего электрического и термического состояния элемента. Иначе говоря, термистор может оставаться еще теплым после работы при высокой температуре, в то время как выделение джоулева тепла в этот момент уже равно нулю.

Термистор характеризуется рядом важных физических, термических и электрических параметров, таких, как:

1. Геометрические размеры, включая размеры баллона и держателей, если они имеются.

2. Свойства материала, из которого изготовлен термистор. Из них важны следующие: удельное сопротивление, температурный коэффициент сопротивления, удельная теплоемкость, плотность и коэффициент теплового расширения.

3. Сопротивление постоянному току R_0 при некоторой исходной температуре T_0. Узаконенная исходная температура составляет 300° К, что близко к «комнатной температуре».

4. Чувствительность по мощности, определяемая как входная мощность в ваттах, необходимая для понижения сопротивления термистора на 1%, или от R_0 до 0,99 R_0.

5. Коэффициент теплоотвода G, определяемый как отношение выделяемой в термисторе тепловой мощности в ваттах к превышению его температуры в градусах над температурой окружающего его пространства в установившемся режиме. Коэффициент теплоотвода является, таким образом, мерой эффективности отвода тепла от термисторного элемента. Коэффициент теплоотвода является важным параметром, поскольку наряду с теплоемкостью он позволяет рассчитать время установления нового стационарного режима при изменении входной мощности.

6. Теплоемкость C, измеряемая количеством тепла в джоулях, запасаемого термисторным элементом, при температуре на 1° С большей температуры окружающей среды. Теплоемкость определяется массой элемента и удельной теплоемкостью его материала.

7. Тепловая постоянная времени τ, представляющая собой время, необходимое для того, чтобы температура элемента достигла величины $(1—1/e)$ температуры установившегося режима при мгновенном изменении входной мощности. Численно тепловая постоянная времени равна $\tau = \frac{C}{G}$. Это соотношение и его смысл рассмотрены в § 3-3-1.

8. Максимальная мощность, рассеиваемая элементом при стабильной работе и без сокращения долговечности.

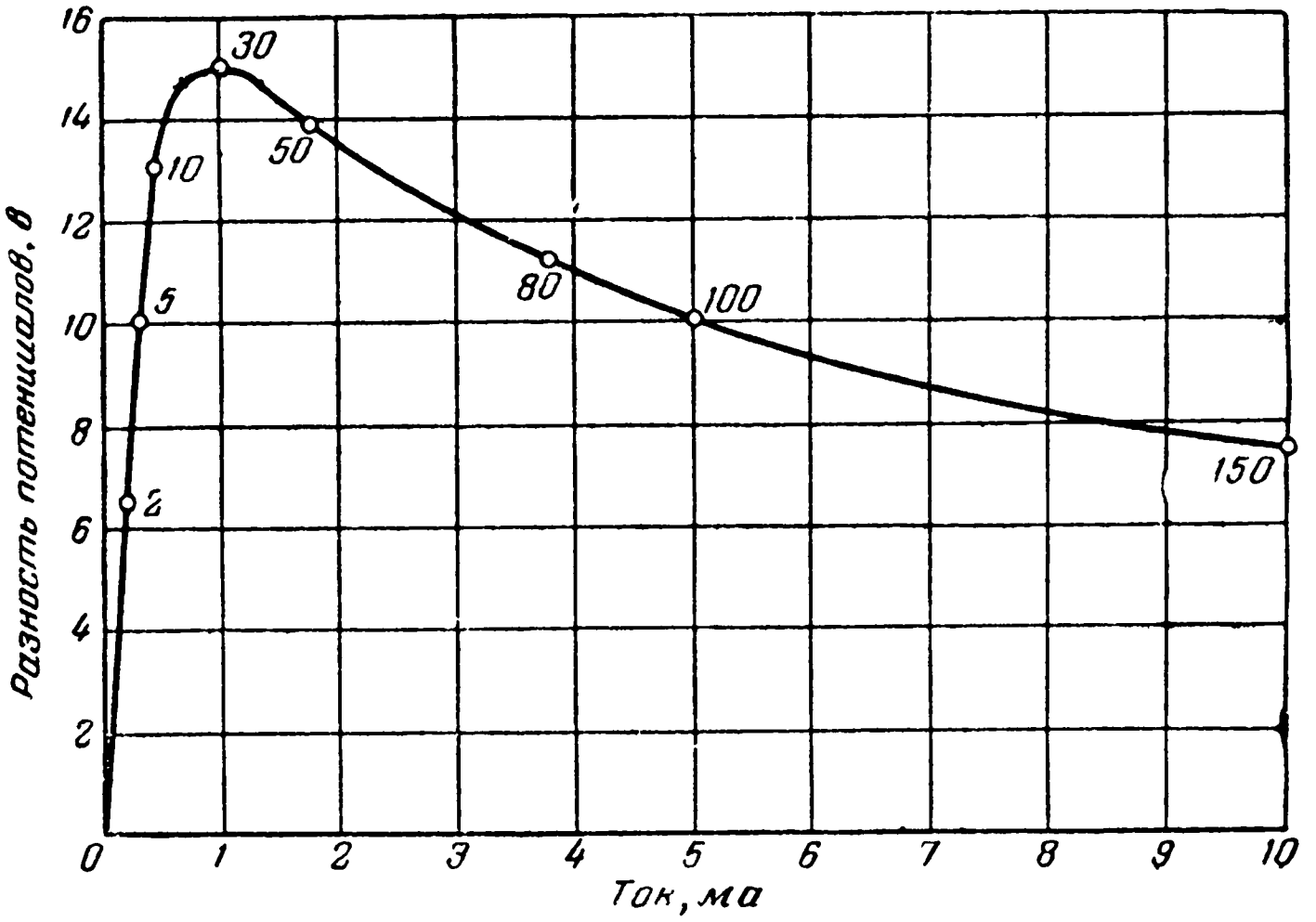


Рис. 3-3. Термистор имеет весьма нелинейную характеристику. Цифры на кривой обозначают превышение температуры термистора в градусах над окружающей средой.

Многие области применения определяются особенностями электрических характеристик термисторов. На рис. 3-3 изображена характеристика термистора с бусинкой диаметром 0,61 *мм*, изготовленной из термисторного материала № 1 * и помещенной в воздухе. При снятии этой характеристики каждый раз после изменения тока делалась достаточная выдержка времени до отсчета разности потенциалов, чтобы температура бусинки устано-

* Термисторный материал № 1 представляет собой смесь окислов марганца и никеля, взятых в такой пропорции, чтобы образовать никель-марганцевый шпинель.

вилась. Поэтому такая кривая называется статической характеристикой.

При малых токах и напряжениях мощность, выделяемая в термисторе, слишком мала, чтобы заметно повысить его температуру; в этой области вблизи начала координат соблюдается закон Ома и характеристика имеет вид прямой линии. Однако при повышенных токах и напряжениях температура бусинки возрастает. При дальнейшем повышении тока сопротивление бусинки понижается, а напряжение на ней уменьшается. Поэтому характеристика отклоняется вправо от продолженной прямолинейной начальной части. При некотором значении тока $I_{макс}$ относительное увеличение тока становится равным вызванному им относительному понижению сопротивления, в результате чего напряжение на термисторе остается постоянным. Этой величине тока соответствует максимум кривой. При еще бóльших токах сопротивление понижается сильнее, чем увеличивается ток, и на характеристике появляется область отрицательного сопротивления. Числа на кривой показывают превышение температуры бусинки над температурой окружающей среды для соответствующих точек характеристики.

Поскольку величины токов и напряжений для различных термисторов охватывают очень широкий диапазон значений, часто бывает удобно строить график для $\lg U$ как функции $\lg I$. На рис. 3-4 приведен в логарифмическом масштабе тот же график, который изображен на рис. 3-3 в линейном масштабе. В логарифмическом масштабе наклон кривой на 45° соответствует постоянному сопротивлению; наклон на минус 45° соответствует постоянной мощности.

Многие практические применения термисторов основаны на одной или нескольких особенностях этой характеристики. Так, чтобы служить простым стабилизатором напряжения, термистор должен работать в режиме, соответствующем вершине кривой. При этом колебания напряжения могут быть ограничены в пределах нескольких процентов при изменениях величины тока в ту или другую сторону на 50%. Часть характеристики с отрицательным сопротивлением используется при применении термисторов в усилителях. При надлежащем режиме термисторы могут работать в экспандерах и компрессорах, когда используется зависимость сопротивления терми-

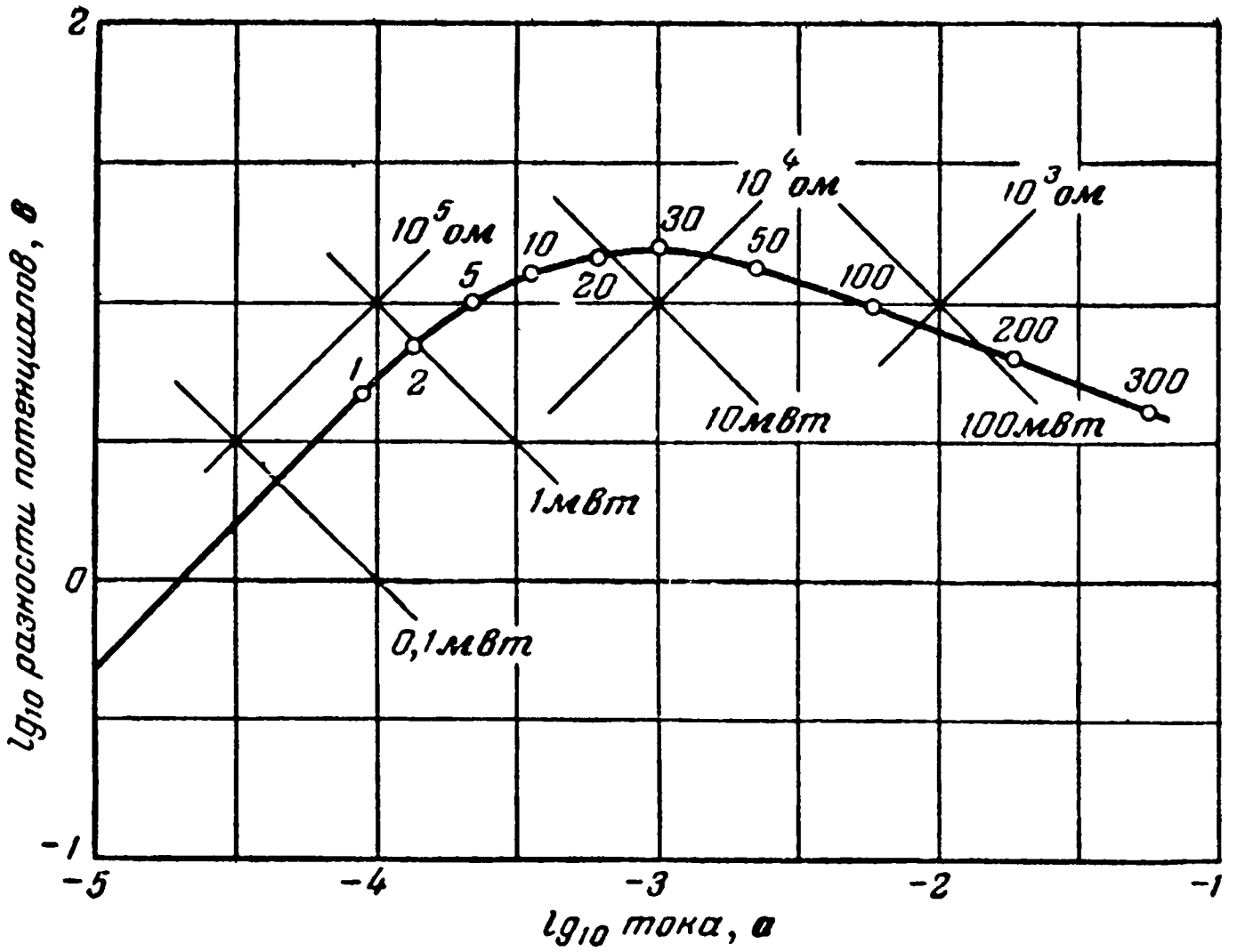


Рис. 3-4. Кривая зависимости рис 3-3, представленная в логарифмических координатах.

стора от подводимой мощности. В следующей главе мы вернемся к этим применениям и обсудим их несколько полнее.

Часто бывает необходимо рассчитать напряжение $U_{макс}$, ток $I_{макс}$, мощность $W_{макс}$, сопротивление $R_{макс}$ или температуру $T_{макс}$, соответствующие максимуму характеристики термистора. Если известны R_0, B, G и T_0, то все указанные выше величины могут быть рассчитаны с помощью следующих формул:

$$T_{макс} = \frac{B}{2}\left(1 - \sqrt{1 - 4T_0/B}\right); \qquad (3\text{-}1)$$

$$W_{макс} = G\left(T_{макс} - T_0\right); \qquad (3\text{-}2)$$

$$R_{макс} = R_0 e^{B(1/T_{макс} - 1/T_0)}; \qquad (3\text{-}3)$$

$$U_{макс} = \left[GR_0\left(T_{макс} - T_0\right) e^{B(1/T_{макс} - 1/T_0)}\right]^{1/2}; \qquad (3\text{-}4)$$

$$I_{макс} = \left[\frac{G}{R_0}\left(T_{макс} - T_0\right) e^{-B(1/T_{макс} - 1/T_0)}\right]^{1/2}. \qquad (3\text{-}5)$$

Для каждого индивидуального термистора логарифмическая характеристика может быть смещена изменением коэффициента теплоотдачи. Это можно осуществить изменением давления газа, окружающего элемент, погружением элемента в теплопроводящую, но электроизолирующую жидкость или любым изменением теплового контакта термисторного элемента и окружающих деталей. На рис. 3-5 приведена логарифмическая харак-

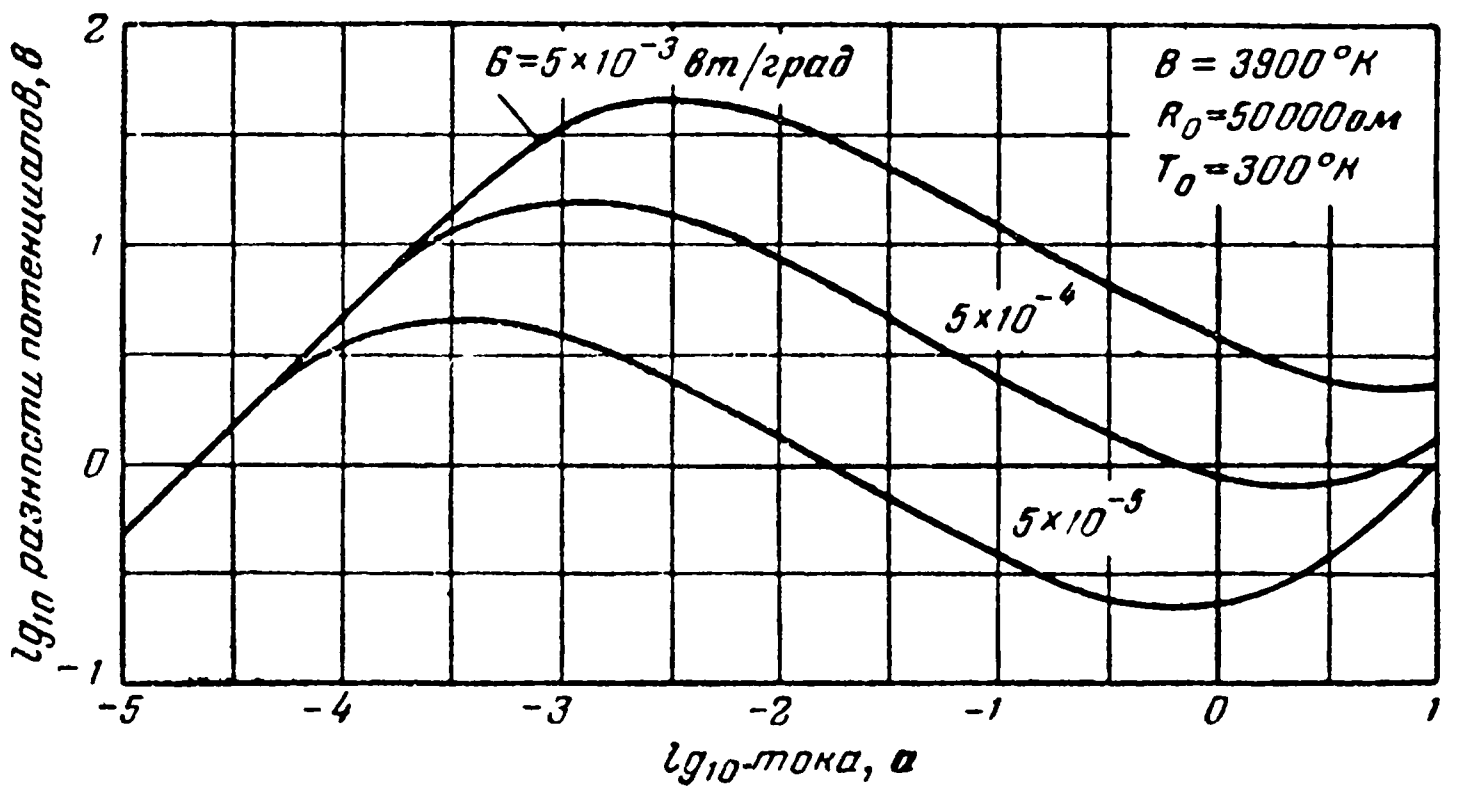


Рис 3-5 Смещение логарифмических характеристик термистора при изменении его коэффициента теплоотдачи.

теристика термистора при трех различных значениях коэффициента теплоотдачи. Обращает на себя внимание то, что эти кривые имеют одинаковую форму и лишь сдвинуты на графике в направлении увеличения мощности. Это происходит потому, что если, например, увеличить отвод тепла от термистора, то понадобится увеличить напряжение на элементе и соответственно, ток через него, чтобы сохранить неизменными его температуру и сопротивление.

Численная величина коэффициента теплоотдачи может быть получена из любой из этих кривых. Для любой точки кривой по отношению U/I находим сопротивление постоянному току R_T. Сопротивление постоянному току R_0 при окружающей температуре T_0 может быть определено из левой части кривой, наклоненной на 45°. Если известны окружающая температура и величина B для данного материала термистора, то можно опреде-

лить температуру термистора в любой точке кривой с помощью уравнения

$$R_T = R_0 e^{B(1/T - 1/T_0)}, \tag{3-6}$$

или

$$T = \frac{1}{\frac{1}{T_0} + \frac{1}{B} \ln \frac{R_T}{R_0}}. \tag{3-7}$$

Для каждой точки выделяемая мощность W может быть определена как произведение UI. Построив по нескольким точкам кривую для W как функцию T, можно получить линию, наклон которой dW/dT дает величину коэффициента теплоотдачи G. Для термисторов, теплоотдача которых происходит главным образом посредством теплопроводности, эта зависимость $W=f(T)$ представляет собой почти прямую линию в довольно широком интервале температур. Если температура термистора превышает температуру окружающего пространства всего на несколько единиц или десятков градусов, отвод тепла действительно происходит главным образом посредством теплопроводности; если обеспечить работу в пределах этого интервала температур, то описанный метод определения G будет вполне точным.

На рис. 3-6 показано смещение логарифмических характеристик термистора, которое происходит вслед-

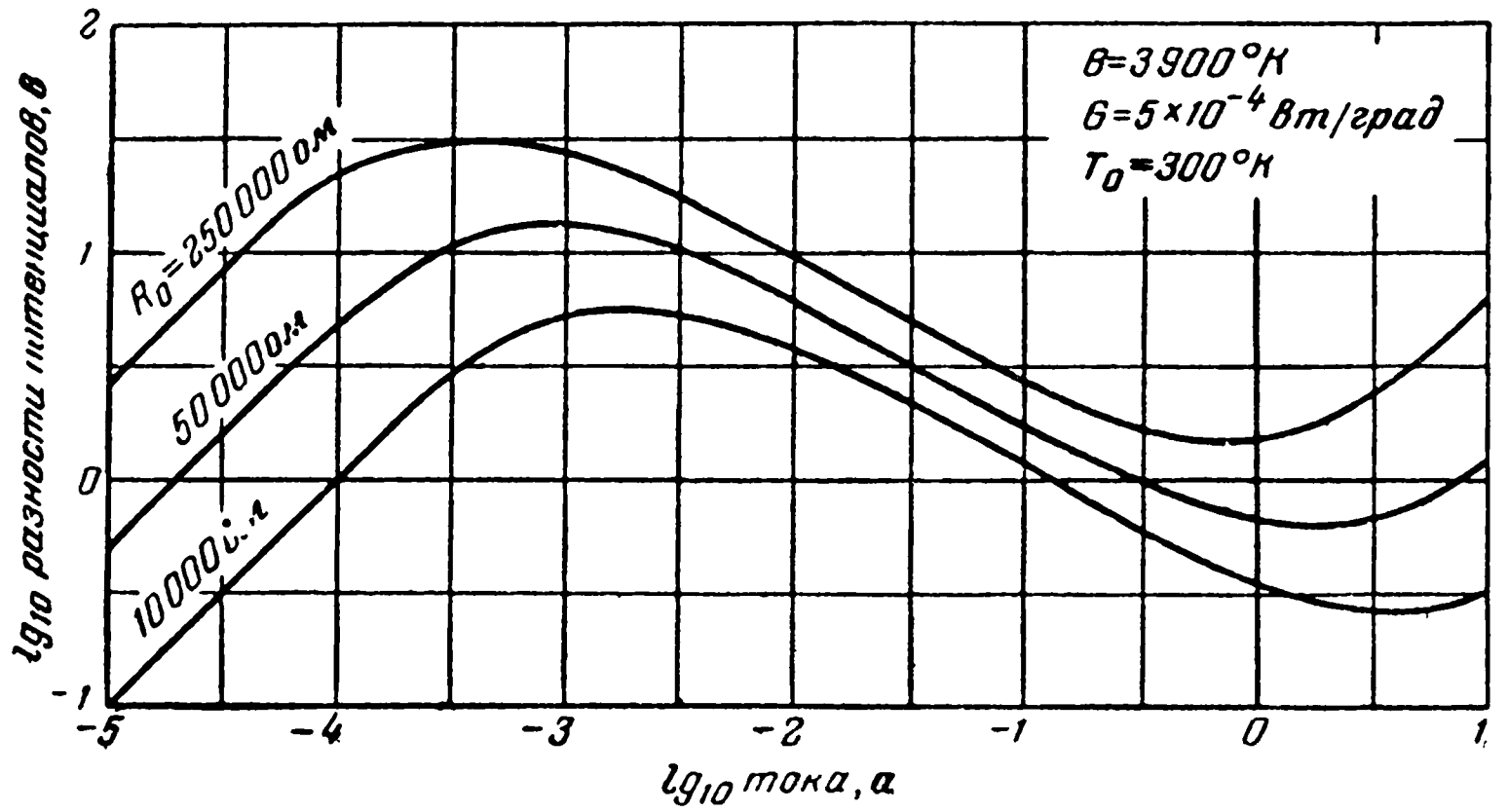


Рис. 3-6. Различие логарифмических характеристик термисторов, имеющих одинаковые значения B и G, но разные R_0.

ствие изменения величины R_0, при B и G постоянных. Достигается это изменением формы полупроводникового элемента при сохранении постоянной величины его наружной поверхности и размеров его держателей. Кривые этого графика можно совместить, сдвигая их в направлении изменения сопротивления.

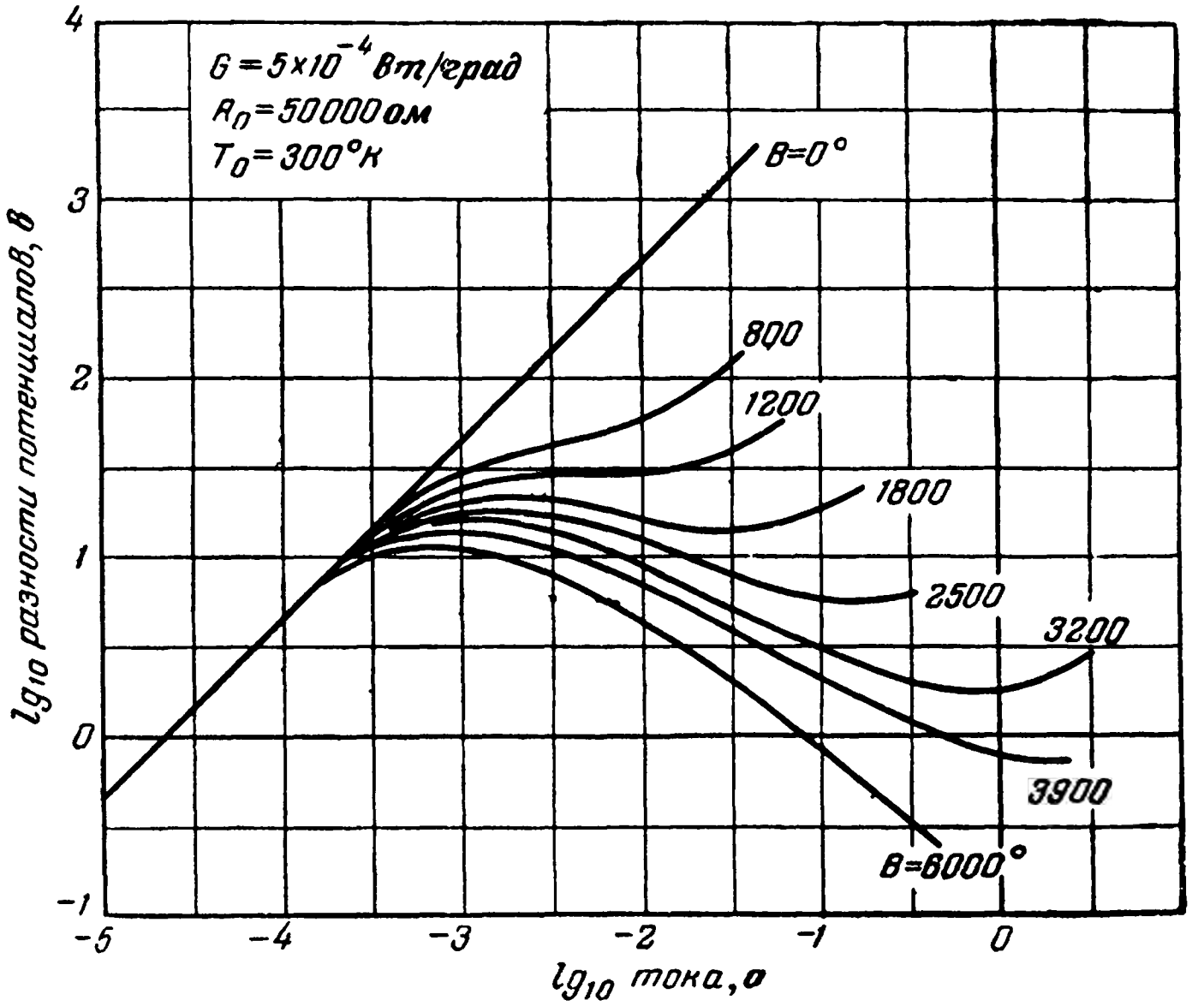


Рис. 3-7 Логарифмические характеристики термисторов, различающихся только величиной B их материала.

Другим переменным параметром, находящимся в распоряжении разработчика, является величина постоянной B материала термистора, которая может быть изменена в довольно широких пределах надлежащим выбором материала.

На рис. 3-7 показано семейство логарифмических характеристик для серии материалов, имеющих различную величину B, из которых были изготовлены термисторные элементы, имеющие одинаковые величины R_0 и G. В то время как показанные на рис. 3-5 и 3-6 характеристики могут быть совмещены простым сдвигом в соответствующем направлении, характеристики на рис. 3-7

совершенно различны. Все кривые имеют одно и то же предельное омическое сопротивление при комнатной температуре, предусмотренное при разработке, но различные предельные сопротивления при больших токах. Для кривой $B=0$ оба эти предельных сопротивления совпадают, и характеристика имеет вид прямой линии. Чем больше значение B, тем сильнее различаются и предельные сопротивления. Для значений B, меньших примерно 1 200° С, характеристики не имеют максимумов.

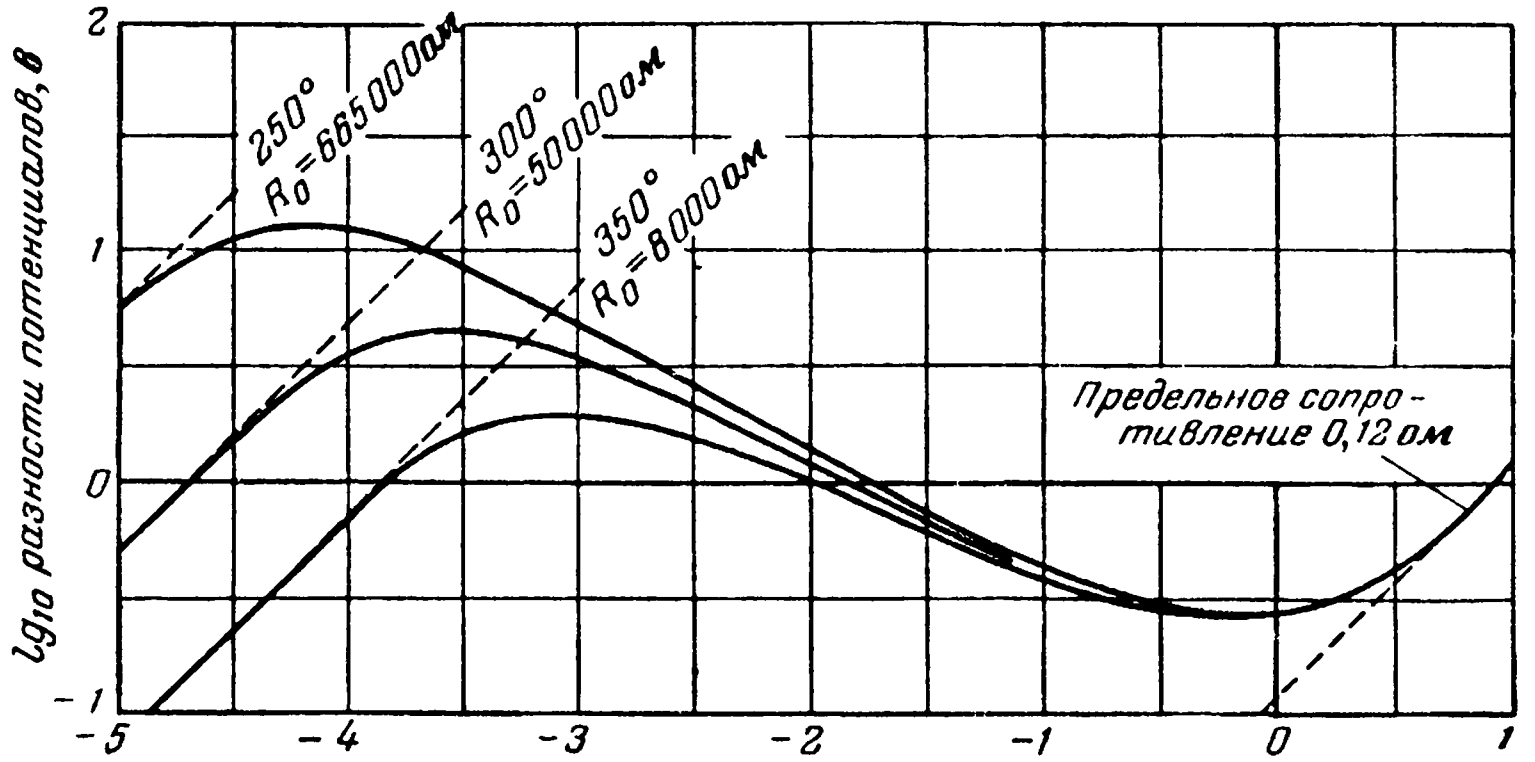


Рис. 3-8 Изменение логарифмических характеристик термистора при изменении температуры окружающей среды $B=3\,900°$ С; $G=5\cdot10^{-5}$ *вт/град*; $R_0=50$ *ком* при температуре окружающей среды 300° С.

Для бóльших значений B существуют максимумы, и чем выше величина B, тем ниже мощность и температура, соответствующие максимуму данной кривой.

Хотя температура окружающей среды не является расчетным параметром термистора, она тем не менее служит важным рабочим параметром с точки зрения использования термистора. Характеристики термистора могут быть изменены вследствие изменения внешней температуры, при которой они снимаются. На рис. 3-8 дана логарифмическая характеристика термистора при комнатной температуре и при температурах выше и ниже указанной на 50° С. Как и следовало ожидать, принципиальная разница между этими характеристиками заключается в положениях наклоненных на 45° асимптот левых частей кривых, определяющих омическое сопро-

тивление элемента при трех значениях температуры окружающей среды. Заметно также, что по мере повышения температуры окружающей среды понижается максимальное напряжение. Последняя особенность используется при применении термисторов в системах пожарной сигнализации, о чем рассказано в следующей главе.

Характеристики на рис. 3-5, 3-6, 3-7 и 3-8 были построены по рассчитанным величинам U для различных значений I, G, B и R_0, принятых за независимые параметры. Поэтому эти кривые несколько идеализированы сравнительно с полученными экспериментально. В любом практическом случае последовательно с сопротивлением термистора включено некоторое сопротивление, хотя бы за счет сопротивления подводящих проводов и контактов Это добавочное сопротивление затрудняет снятие характеристик полупроводниковых элементов при высокой температуре, когда их собственное сопротивление низко. Более того, при температуре термистора, превышающей окружающую более чем примерно на 100° С, на сцену выступают потери вследствие излучения тепла, которые играют тем бóльшую роль, чем выше температура. Поскольку потери на излучение подчиняются закону

$$\textit{потери на излучение} \approx T^4 - T_0^4, \qquad (3\text{-}8)$$

т е. нелинейно зависят от температуры T, коэффициент теплоотдачи оказывается не постоянной величиной, как это было допущено при расчетах, а зависит от входной мощности. Наконец, в некоторых случаях асимптота характеристики со стороны малых сопротивлений может соответствовать температурам, превышающим ту, при которой некоторые части термистора выходят из строя. В таких случаях крайняя правая часть характеристики не может быть снята экспериментально.

3-3-1. Свойства термисторов, зависящие от времени

До сих пор рассмотрение ограничивалось условиями установившегося режима, когда подводимая к термистору мощность равна выделяемой при постоянной температуре. Однако во многих случаях необходимо бывает рассмотреть переходные процессы, при которых температура и все величины, являющиеся функциями температуры, меняются во времени.

Рассматривая установившийся и переходные режимы термисторов, полезно пользоваться аналогиями. Можно «ухватить смысл» поведения незнакомой системы, подобрав аналогичную знакомую систему и исследовав ее поведение. Термистор является тепловым аналогом электрической цепи, состоящей из последовательно соединен-

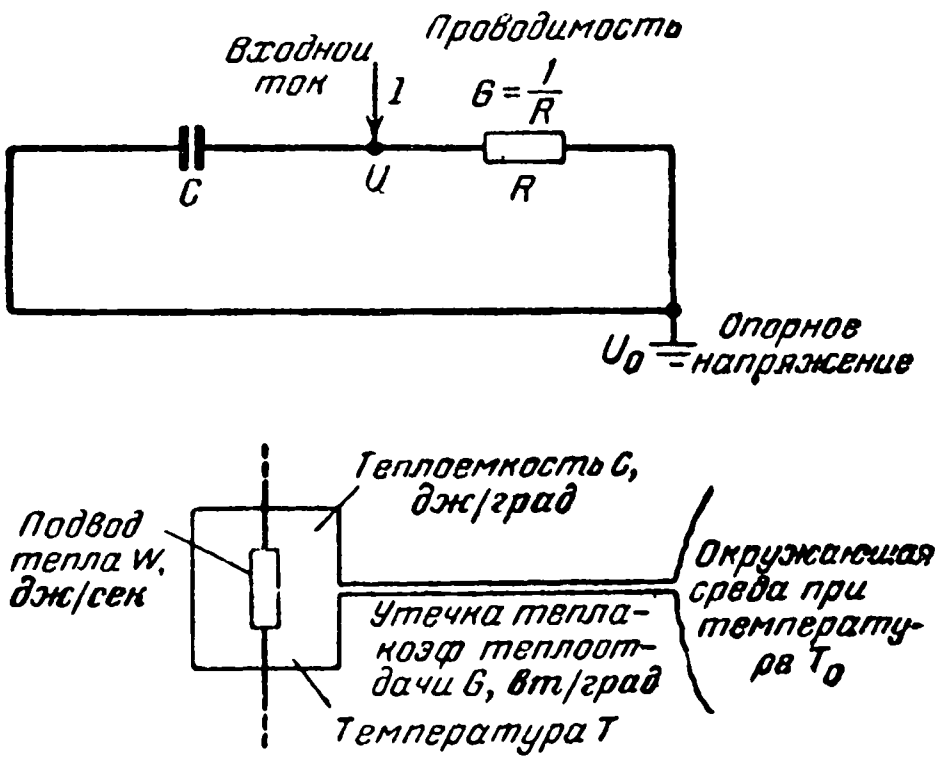


Рис. 3-9. Термистор (внизу) представляет собой тепловую систему, аналогичную электрической цепи, показанной вверху. Уравнения, описывающие тепловые и электрические процессы в этой системе и цепи, одинаковы. Установившийся ток при входном токе I: $U-U_0=I/G$; спад после выключения тока на входе: $U-U_0=I/Ge^{-(G/C)t}$, установившийся режим при потоке тепла W. $T-T_0=W/G$; спад после прекращения тепловыделения $T-T_0=W/Ge^{-(G/C)t}$.

ных конденсатора и сопротивления. На рис. 3-9 показаны тепловая цепь термистора и ее электрический аналог с указанием соответствующих частей. Для облегчения этого сравнения в табл. 3-1 приведен список аналогичных понятий в «тепловых» и «электрических» терминах.

Термоэлектрическая модель может быть использована следующим образом. Обращаясь к верхней схеме на рис. 3-9, каждый, знакомый с элементарной теорией электрических цепей, может сразу написать для постоянного установившегося тока I:

$$U-U_0=\frac{I}{G}\,. \tag{3-9}$$

Таблица 3-1

Аналогичные электрические и тепловые понятия

Понятие	Единицы	
	электрические	тепловые
Напряжение	Вольты	Градусы
Количество	Кулоны	Тепловые джоули
Поток	Кулоны в секунду = =амперы	Тепловые джоули в секунду=тепловые ватты
Проводимость	G, ампер на вольт= =мо	G, коэффициент теплоотдачи=тепловые джоули в секунду на градус
Емкость	C, кулон на вольт= =фарады	C, теплоемкость—тепловые джоули на градус
Постоянная времени	$\tau=RC=C/G$ *сек*	$\tau=C/G$ *сек*

Аналогично для термистора, получающего W *вт* установившейся входной тепловой мощности при выделении W *вт* электрической мощности в элементе, можно написать аналог уравнения (3-9), подставив вместо электрических тепловые обозначения:

$$T - T_0 = \frac{W}{G}, \qquad (3\text{-}10)$$

что справедливо, как мы уже знаем, из определения G.

Чтобы определить, как спадает температура термистора после мгновенного выключения мощности нагрева, можно вместо выполнения специальных вычислений просто выяснить, что происходит с напряжением в электрической модели, когда входной ток мгновенно прерывается в момент времени $t=0$. Это напряжение равно установившемуся напряжению в момент разрыва тока, помноженному на знакомый множитель, выражающий экспоненциальный спад:

$$U - U_0 = \frac{I}{G} e^{-t/RC} = \frac{I}{G} e^{-(G/C)t} = \frac{I}{G} e^{-t/\tau}, \qquad (3\text{-}11)$$

где

$$\tau = RC = \frac{C}{G}.$$

Соответственно для термистора спад температуры после выключения подогрева выражается так:

$$T - T_0 = \frac{W}{G} e^{-t/\tau}, \qquad (3\text{-}12)$$

где

$$\tau = \frac{C}{G}.$$

Уравнение (3-12) показывает, что нагретый термисторный элемент должен остывать экспоненциально с постоянной времени, равной C/G. Поэтому график логарифма $(T—T_0)$, построенный как функция времени,

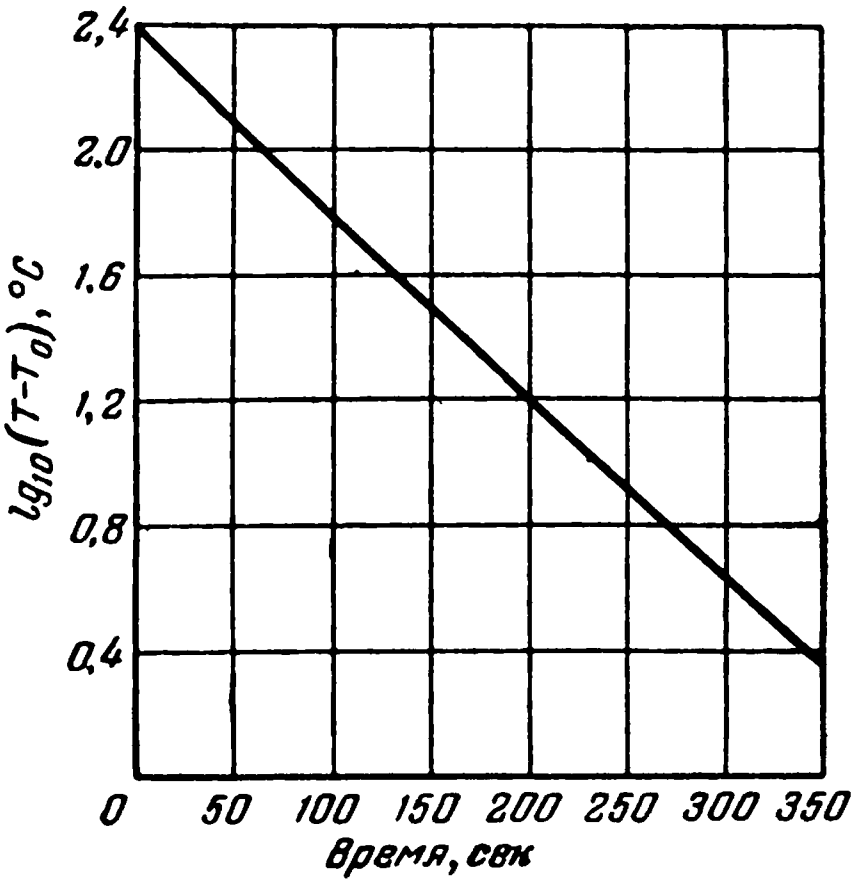


Рис 3-10. Кривая охлаждения большого термистора, предварительно нагретого до 250° С.

должен иметь вид прямой линии. На рис. 3-10 показана такая кривая охлаждения, найденная экспериментально для большого термистора, предварительно нагретого примерно на 250° С выше комнатной температуры. Измерения производились при пропускании через термистор измерительного тока достаточно слабого, чтобы не влиять на процесс его охлаждения. Кривая действительно очень близка к прямой линии; имеющиеся отклонения объясняются, возможно, влиянием излучения тепла

и небольшого изменения теплоемкости C при изменении температуры.

Если тот же термисторный элемент, первоначально холодный внезапно поместить в нагретую среду, он будет нагреваться соответственно кривой, имеющей тот же логарифмический наклон, что и на рис. 3-10, но с положительным знаком. Однако если термистор, первоначально имевший температуру окружающей среды, нагревается после подачи напряжения на его зажимы, то кривая нагрева не подчиняется простому закону нарастания со временем, поскольку скорость подъема температуры пропорциональна входной мощности, зависящей от мгновенного сопротивления элемента, которое в свою очередь зависит от температуры. Этот процесс выражается следующим дифференциальным уравнением:

$$\frac{U^2}{R_0} e^{-B(1/T-1/T_0)}\, dt = G(T - T_0)\, dt + C dT. \qquad (3\text{-}13)$$

Левая часть его представляет собой джоулево тепло, выделяемое электрическим током в элементе за вре-

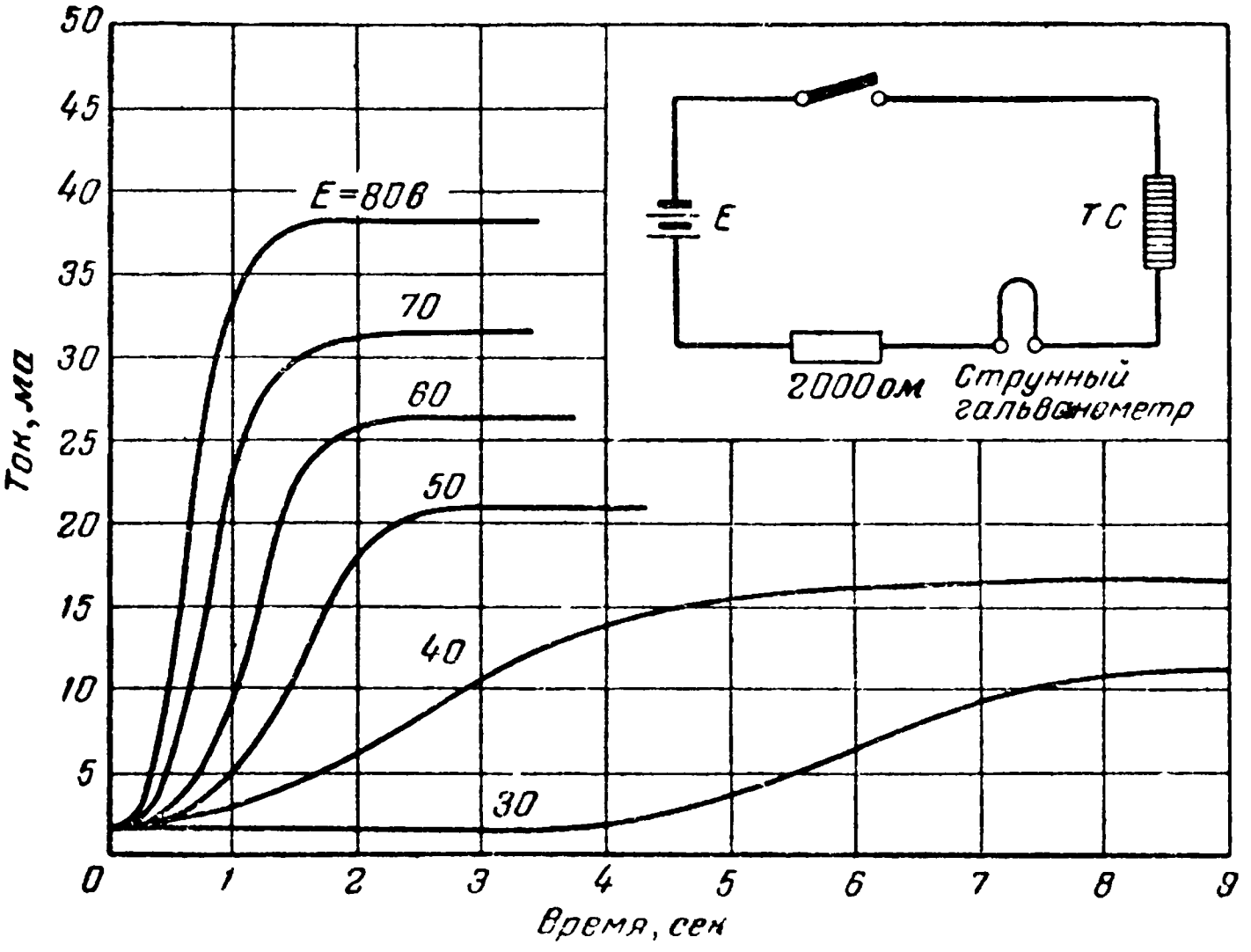


Рис. 3-11. Семейство кривых нарастания тока при разных напряжениях питания цепи задержки (верхняя часть графика).

мя *dt*. Первый член правой части выражает ту часть подводимого тепла, которая необходима для компенсации теплоотвода за то же время, тогда как второй член правой части есть оставшаяся часть тепла, которая вызывает нагрев элемента на величину *dT*. Аналитическое решение этого уравнения затруднительно [Л. 2].

На рис. 3-11 дано семейство кривых для почти аналогичной электрической модели. Время отсчитывается от момента включения тока. Каждая кривая дана для определенной величины питающего напряжения. В любой практической цепи всегда должно иметься некоторое внешнее защитное сопротивление; это сопротивление усложняет явление, вызывая падение напряжения на термисторе по мере возрастания тока, затрудняя еще больше аналитическое решение уравнения (3-13).

Кривые на рис. 3-11 являются основой для использования термисторов как элементов задержки времени. Для этого в их цепь можно включить реле, отрегулированное на срабатывание в момент, когда ток достигает определенной величины. Можно сильно изменять величины задержек, варьируя напряжение питания и подбирая величины C, G и τ в зависимости от надлежащего выбора материалов, размеров и вида теплоотводящей среды. С помощью таких схем на термисторах можно осуществить задержки во всем диапазоне длительностей от миллисекунд до примерно 10 *мин*.

3-3-2. Быстродействующие термисторы

До сих пор мы рассматривали быстроту реакции термисторов на изменение условий. Теперь мы рассмотрим устройство и действие термисторов, рассчитанных на работу при высокой частоте изменения входной мощности и флуктуаций температуры. Такие приборы называются быстродействующими (малоинерционными) термисторами. Они представляют собой обычные термисторы, в конструкции которых особое внимание уделено выбору геометрии полупроводникового элемента и его электрических контактов для снижения теплоемкости и повышения теплоотвода. В обоих случаях (порознь или одновременно) происходит уменьшение постоянной времени.

Несколько возможных схем решения этой задачи приведено на рис. 3-12. На схеме *а* показан кристалл полупроводника, прижатый к металлической поверхно-

сти. Из-за явления распределенного сопротивления, разобранного в гл. 6, большая часть эффективного электрического сопротивления термистора оказывается сосредоточенный в малом объеме полупроводника, непосредственно примыкающем к точке контакта между кристаллом и металлической пластинкой. Этот объем вследствие его малости имеет малую теплоемкость и быстро отдает тепло ввиду хорошего теплового контакта с металлической пластиной и с основной массой полупроводника. На схемах *б* и *в* показан слой полупроводника

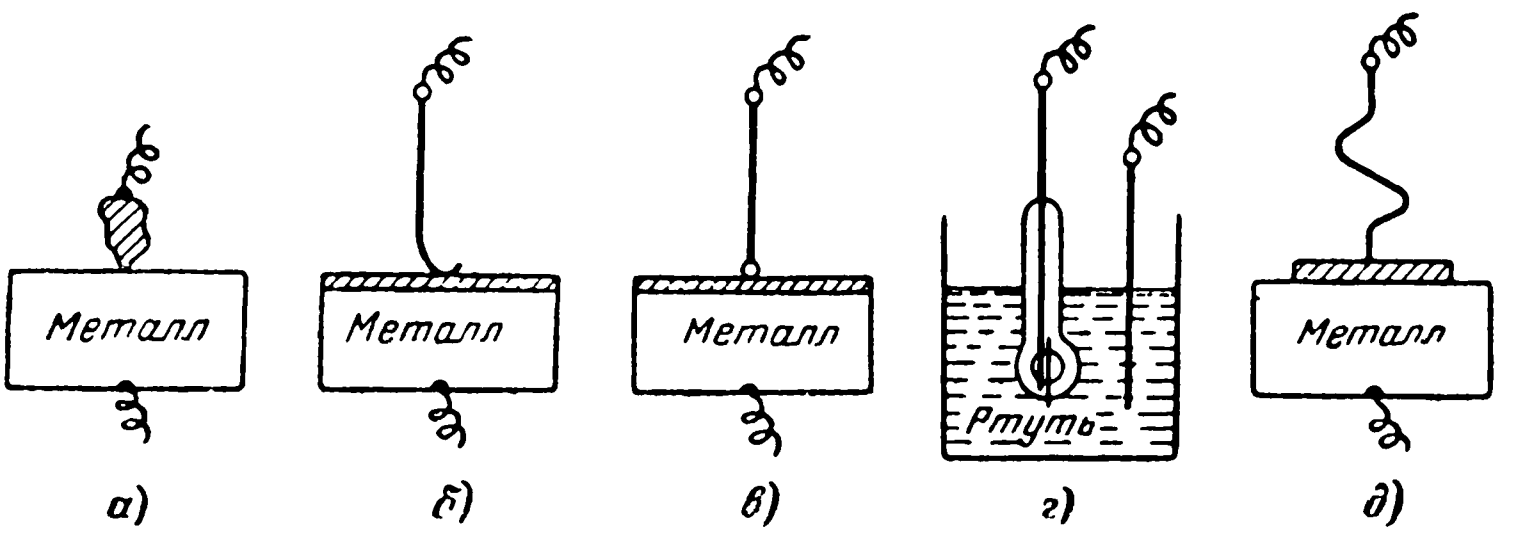


Рис. 3-12 Схемы малоинерционных термисторов; теплоемкость мала, а коэффициент теплоотдачи велик.

в плотном контакте с металлической подкладкой. Со свободной поверхностью полупроводника на схеме *в* соприкасается маленький шарик, образуя точечный контакт, тогда как на схеме *б* электродом служит согнутая пружинящая проволочка, прикасающаяся к поверхности полупроводника. В обоих вариантах эффективные объемы полупроводника малы, а хороший теплоотвод обеспечивается геометрией слоя. На схеме *г* использован обычный бусинковый термистор. Теплоемкость такого термистора всегда одинакова при одинаковых размерах, форме и материале, но коэффициент теплоотдачи значительно увеличивается, если поместить его в тонкую стеклянную оболочку и погрузить в ртуть или впаять сплавом Вуда в металлический блок. На схеме *д* показан германиевый малоинерционный термистор. Точечный контакт образуется между позолоченной вольфрамовой проволокой и сплошным германием, сильно легированным сурьмой или алюминием. Такой германий имеет проводимость *p*-типа, но не выпрямляет в точечном контакте.

В большинстве случаев использования малоинерционных термисторов работа происходит на спадающей части статической характеристики при токе подогрева, превышающем тот, который соответствует максимуму напряжения на рис. 3-3. Предположим, что режим термистора отмечен точкой P на рис. 3-13 *. Если на постоянный ток

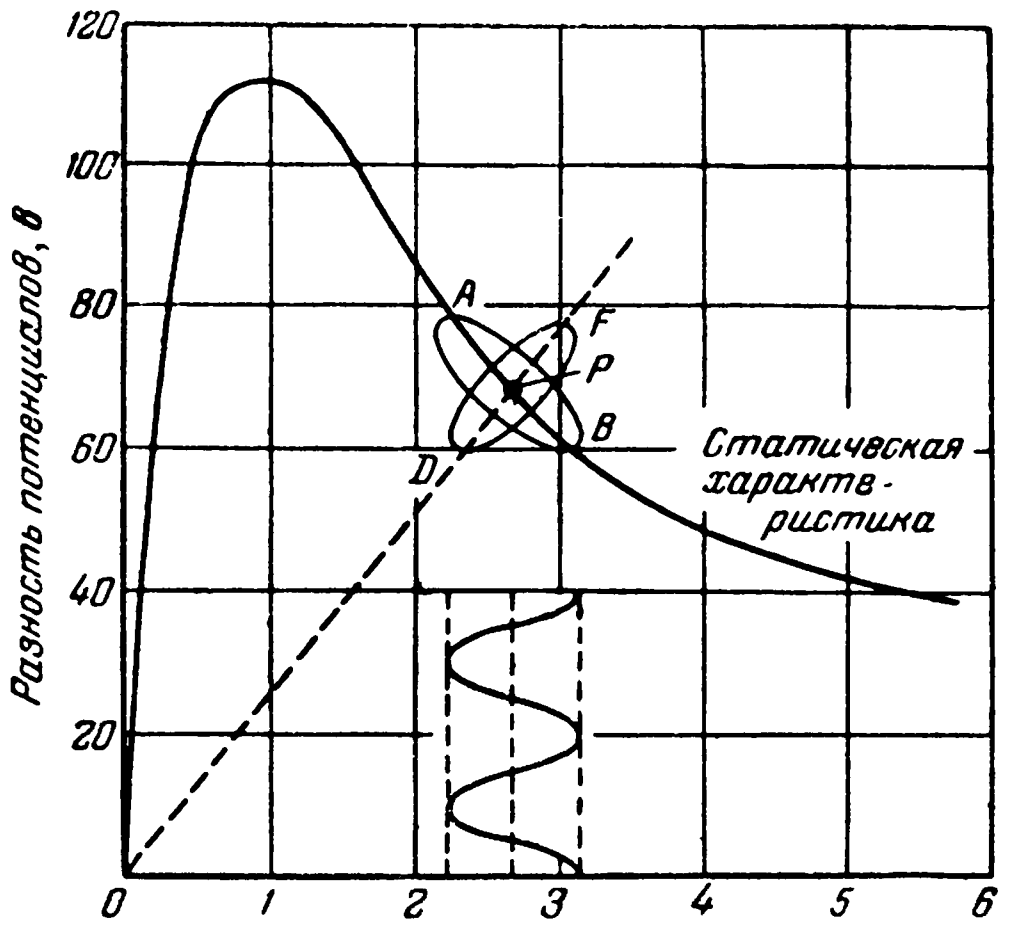


Рис 3-13. Поведение термистора при небольшой амплитуде переменных сигналов разной частоты. A—B — низкая частота, D—F — высокая частота

подогрева накладываются колебания небольшой частоты и амплитуды, то рабочая точка мгновенного режима будет совершать колебания в обе стороны вдоль линии APB. В этих условиях сопротивление прибора переменному току dU/dI отрицательно и численно выражается наклоном характеристики в точке P. Однако чем выше частота колебаний тока, тем больше мгновенное значение температуры термистора отстает от мгновенной мощности из-за его тепловой инерции. Это отставание фазы заставляет рабочую точку мгновенного режима описывать эллипс, большая ось которого при низких ча-

* Краткое пояснение способа установки режима и метода нагрузочной линии для решения простых задач по анализу цепей см § 3-6.

стотах колебаний лежит почти вдоль линии *APB*, а при высоких частотах — почти вдоль линии сопротивления постоянному току *DPF*. Таким образом, сопротивление термистора на низких частотах отрицательно; оно остается отрицательным, но меньшим по абсолютной величине при увеличении частоты, проходит через нуль и становится положительным, все бóльшим по величине на еще более высоких частотах, приближаясь в пределе к линии *DPF*. Частота, при которой сопротивление переменному току проходит через нуль, называется критической частотой термистора. В табл. 3-2 приведены критические частоты, характерные для различных типов малоинерционных термисторов, показанных на рис. 3-12.

Таблица 3-2

Критические частоты различных типов малоинерционных термисторов

Тип	Конструкция	Критическая частота, *кгц*
а	Кристалл бора на металлической поверхности	2
б	Пружинный проволочный контакт на тонком слое U_3O_8	7
в	Сферический металлический контакт на тонком слое U_3O_8	4
г	Бусинковый термистор в ртути	0,2
д	Точечный контакт на германии	100

Хотя малоинерционные термисторы до сих пор не использовались в широких промышленных масштабах, был предложен или находится в разработке ряд интересных и полезных применений их. Отрицательное сопротивление переменному току малоинерционных термисторов делает возможным использование таких приборов в качестве активных элементов цепей в генераторах, мультивибраторах и генераторах колебаний специальной формы. Осуществимость таких устройств была показана в лабораторных масштабах. Отрицательное сопротивле-

ние также может быть использовано в промежуточном усилителе для компенсации положительного сопротивления участков цепи. При использовании термисторов в усилителях с отрицательным сопротивлением германиевые малоинерционные термисторы могут дать в полосе звуковых частот усиление до 13 *дб*. Тепловая инерция в термисторах, работающих на частотах ниже критической, электрически соответствует индуктивности в экви-

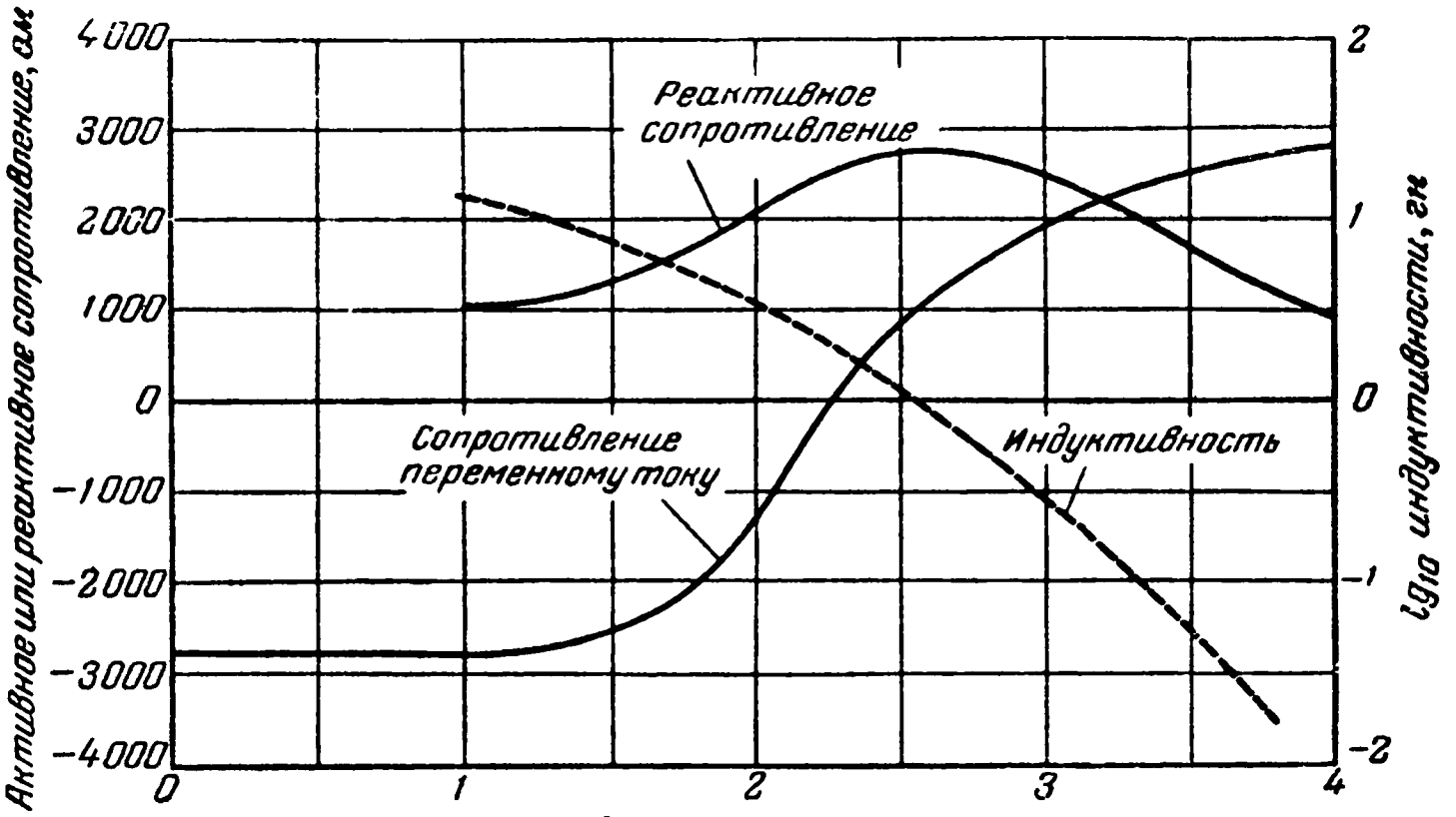


Рис 3-14. Сопротивление термистора при малых амплитудах переменного сигнала имеет активную и реактивную составляющие, зависящие от частоты.

валентной схеме термистора. Величина этой индуктивности зависит от частоты таким образом, что реактивное сопротивление достигает максимума вблизи критической частоты. Зависимости сопротивления переменному току, индуктивности и реактивного сопротивления от частоты показаны на рис. 3-14 для бусинкового термистора диаметром 0,25 *мм* в стеклянной оболочке, погруженного в ртуть. Индуктивность, проявляемая малоинерционными термисторами, вызвала предложение использовать эти приборы как нагрузочные индуктивности в линиях передачи.

Малоинерционный термистор может быть использован как модулятор. Для несущей частоты, лежащей намного выше критической, эффективное сопротивление термистора положительно. С другой стороны, модулирующий сигнал, имеющий частоту ниже критической,

модулирует сопротивление, через которое проходит ток несущей частоты. Поскольку сопротивление термистора на частоте модулирующего сигнала отрицательно, с помощью такого модулятора может быть произведено усиление модулирующего сигнала.

В будущем представляется перспективным использовать германиевые термисторы ввиду сравнительно высокой их критической частоты и относительной стабильности во времени. Много таких термисторов непрерывно работало в течение более года в генераторах, усилителях, модуляторах и переключателях, в некоторых случаях при выделении в них мощности порядка 1 *вт* без какого-либо заметного изменения характеристик.

3-4. ПОДОГРЕВНЫЕ ТЕРМИСТОРЫ

Все рассмотренные до сих пор термисторы являлись самоподогревными двухполюсниками, в которых мощность, предназначенная для поднятия температуры элемента, вводилась через те же зажимы, которые служили для измерения таких зависимых параметров, как, например, сопротивление или падение напряжения. Иногда отделение управляемой величины от управляющей осуществляется с помощью частотного разделения. Так, в высокочастотном измерителе мощности входная мощность подводится в виде тока высокой частоты, а отсчет ее величины по сопротивлению термистора производится на постоянном токе. Однако для некоторых случаев применения желательно электрическое разделение управляемой цепи от управляющей. В этих случаях могут быть использованы подогревные термисторы.

В подогревном термисторе полупроводниковый элемент обычно представляет собой бусинку, окруженную тонкой подогревательной катушкой. Бусинка и катушка закреплены в стеклянной оболочке в определенном положении друг относительно друга, так что получается четырехполюсник с полностью электрически разделенными подогревателем и бусинкой. Пропускание подогревающего тока через катушку позволяет получить управляемый сдвиг характеристики полупроводниковой бусинки. На рис. 3-15 приведено семейство таких характеристик для бусинки диаметром 0,38 *мм* из термисторного материала № 2, заключенной в стеклянный цилиндр длиной 3,8 *мм* и даметром 0,76 *мм*. На этот цилиндр намотана

и залита стеклом катушка из тонкой нихромовой проволоки сопротивлением 100 *ом*. Для разных значений тока подогрева получились различные характеристики. Влияние повышения тока подогрева состоит в повышении местной температуры окружающего пространства, от кото-

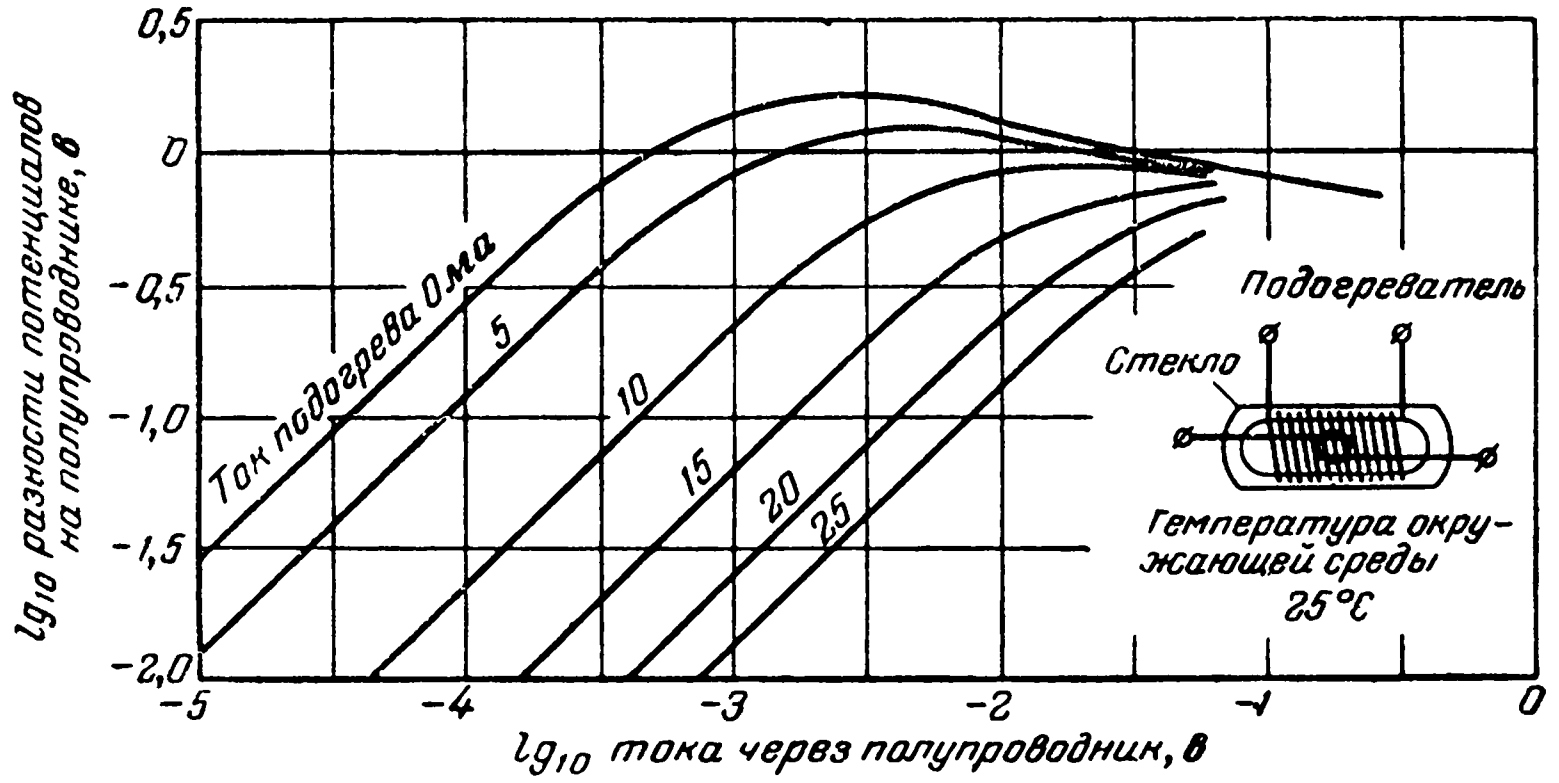


Рис 3-15 Зависимость статических характеристик подогревного термистора от тока в подогревателе

рой зависит поведение термисторного элемента; относительные положения характеристик на рис. 3-15 согласуются с поведением термистора при различных окружающих температурах, иллюстрированным рис. 3-8.

Подогревные термисторы используются для автоматической регулировки усиления. Это применение их описано в гл. 4.

3-5. СТАРЕНИЕ И ШУМЫ

Важным свойством термистора является характеристика его старения. Она представляет собой зависимость электрических параметров термистора от времени его хранения или фактической работы. Для того чтобы обеспечить максимальную стабильность рабочих параметров, при разработке конструкции термистора следует:

1) выбирать только такие полупроводниковые материалы, которые имеют чисто электронную проводимость, химически стойки против воздействия атмосферы при повышенной температуре и нечувствительны к загрязнениям, могущим иметь место в процессе изготовления или при эксплуатации;

2) применять такую обработку материала, чтобы распределение в нем примесей, влияющих на параметры прибора, было вполне установившимся, а если это невозможно, то чтобы это установление при рабочей температуре происходило чрезвычайно медленно;

3) делать контакты плотными, надежно держащимися, с низким переходным сопротивлением, которое должно быть как можно ближе к чисто активному; материал контактов должен соответствовать материалу термистора по коэффициенту теплового расширения;

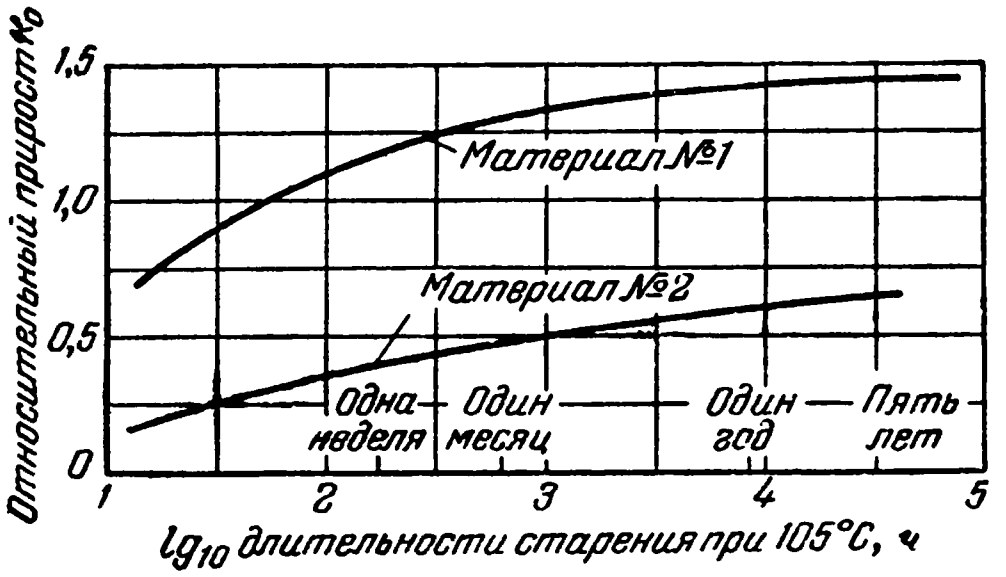


Рис 3-16 Стабильность термисторов при повышенной температуре

4) если необходимо — выбрать подходящее защитное покрытие для элемента; хорошим покрытием может быть заливка стеклом или керамическая гильза; защиту может создать также газовая или жидкая среда, введенная во внешнюю оболочку;

5) применять искусственное старение (тренировку) элемента в течение нескольких дней или недель при температуре не ниже максимальной, ожидаемой при эксплуатации.

При соблюдении этих предосторожностей можно получить характеристики, стабильные с точностью до долей процента в течение многих лет после первоначального искусственного старения. На рис. 3-16 показаны характеристики старения по сопротивлению для двух типичных образцов из двух различных материалов. Как видно, в основном старение происходит в первое время работы изделия. Из рис. 3-16 видно также, что после искусственного старения в течение 1 мес. при 105° С в последующие 5 лет сопротивление при этой температуре изменялось всего примерно на 0,2%. В термисторном термо-

метре это изменение соответствует неточности измерения температуры примерно 0,05° С. Термисторы, смонтированные в откачанных стеклянных баллонах, стареют еще медленнее. Однако в большинстве случаев применения термисторов последствиями старения можно вовсе пренебречь и применять их без всякого искусственного старения.

Специальные испытания на старение при высокой температуре, проведенные на стержневых термисторах, изготовленных методом выдавливания, показали, что в течение месяца при температуре 300° С сопротивление типичных образцов меняется на величину 0,5—1,5% начального значения. Испытывалось воздействие колебаний температуры от +300 до —75° С на подобные термисторы в течение 700 циклов длительностью по 0,5 *ч*. Сопротивления типичных образцов изменились в течение испытания менее чем на 1%.

В некоторых случаях термисторы используются для индикации очень небольших изменений температуры или очень малой разности высокочастотных мощностей. Такие малые приросты измеряемых величин вызывают малые изменения сопротивления полупроводникового элемента. Если сопротивление при этом беспорядочно изменяется из-за тепловых шумов, дробового эффекта в контактах или шума тока в элементе или на его поверхности, то этим ставится минимальный предел изменения измеряемой величины, ниже которого это изменение не может быть отмечено. Для получения высокой чувствительности важно свести шумы к минимуму. Тепловой шум, очевидно, неустраним, но, к счастью, он редко является основной составляющей общих шумов. Контактные шумы и шумы от прохождения тока можно уменьшить путем тщательного конструирования и изготовления изделий. Наиболее важно добиваться однородного спекания материала. Контакты должны быть плотными, надежно закрепленными механически и не обладать электрическим вентильным действием (односторонней проводимостью). Уровень шумов повышают механические напряжения на поверхности полупроводника, возникающие от плохого согласования коэффициентов теплового расширения полупроводника и контактов; эти коэффициенты необходимо согласовать, как и для оболочки, в которую заделан элемент.

3-6. ЗАМЕЧАНИЯ О ВЫБОРЕ РЕЖИМА И ГРАФИЧЕСКОМ РАСЧЕТЕ ЦЕПЕЙ

Решение задач, включающих расчет токов и напряжений в цепях, содержащих нелинейный элемент, включенных последовательно с омической нагрузкой и источником питания, может быть значительно упрощено применением графического метода, известного под названием метода нагрузочной линии. Пусть в цепи, показанной на рис. 3-17, идет ток $I_ц$, величину которого мы желаем определить. Обозначим через U_B напряжение источника

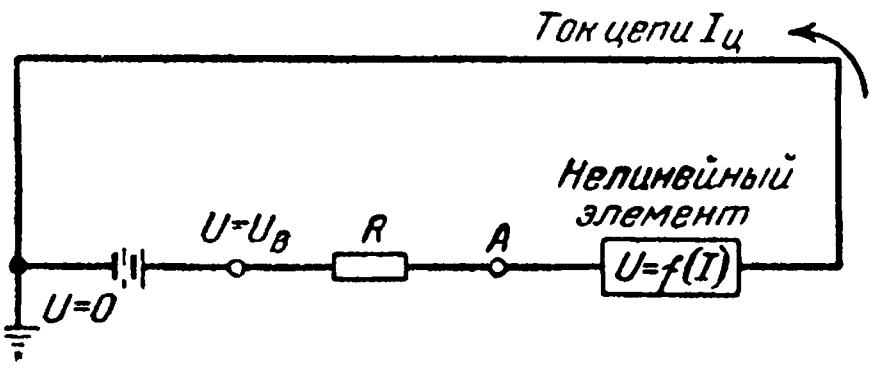


Рис. 3-17. Напряжение в точке A определяется двумя уравнениями $U=U_B—I_цR$ и $U=f(I_ц)$, из совместного решения которых определяется не только это напряжение, но и величина тока в цепи

питания, R — величину сопротивления нагрузки, и пусть характеристика нелинейного элемента описывается функцией $U=f(I)$, график которой задан.

Тогда напряжение U_A в точке A цепи находится аналитически из уравнения

$$U_A=U_B-I_цR. \tag{3-14}$$

Напряжение в этой точке также выражается через

$$U_A=f(I_ц). \tag{3-15}$$

U_A находится путем совместного решения этих двух уравнений. Решение может быть выполнено графически, как показано на рис. 3-18, построением характеристики нелинейного элемента $U=f(I)$ и наложением на нее прямой линии $U=U_B—IR$. Эта прямая называется нагрузочной. Она имеет точку пересечения с осью потенциалов U_B, точку пересечения с осью тока U_B/R и отрицатель-

ный наклон R. Точка ее пересечения с характеристикой нелинейного элемента определяет рабочую точку цепи при $U=U_A$ и $I=I_{ц}$.

Теперь предположим, что напряжение возрастает от U_B до U'_B. При этом нагрузочная линия перемещается в положение, обозначенное на рис. 3-18 пунктиром, и рабочая точка перемещается по характеристике из точки P в точку P'. Обратите внимание, что это смещение, обусловленное повышением напряжения питания, приводит

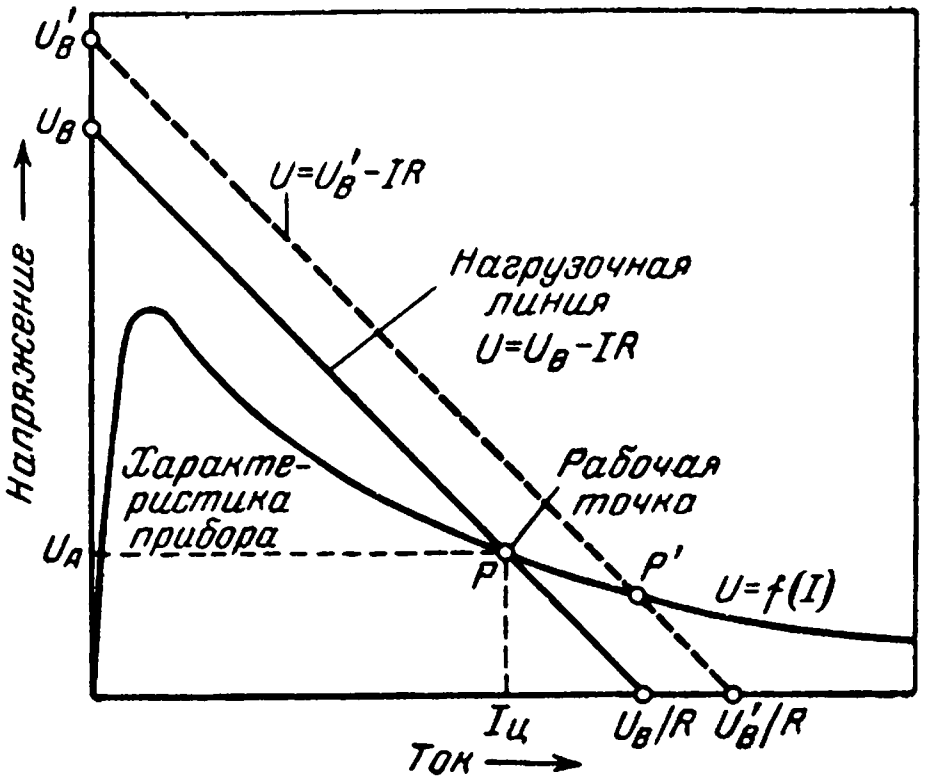


Рис. 3-18 Точка пересечения нагрузочной линии с характеристикой нелинейного элемента определяет ток в цепи и падение напряжения на элементе.

к увеличению тока в цепи и *уменьшению* падения напряжения на линейном элементе. Таким образом, нелинейный элемент проявляет отрицательное сопротивление на участке характеристики от P до P', и именно оно используется при работе в режиме усиления и генерирования. Заметьте, что это отрицательное сопротивление проявляется только при *изменении* положения рабочей точки на участке характеристики, имеющем отрицательный наклон dU/dI. Отрицательное сопротивление проявляется поэтому только для переменного тока. Для постоянного тока нелинейный элемент имеет положительное сопротивление U/I, величина которого задается наклоном прямой, соединяющей точку P с началом координат.

Если желательно выбрать режим нелинейного элемента по постоянному току, соответствующий некоторой вы-

бранной рабочей точке на его характеристике, то нужно выбрать такие величины напряжения источника питания и последовательного сопротивления нагрузки, чтобы нагрузочная прямая, соответствующая этим значениям, пересекала бы характеристику элемента в желательной рабо-

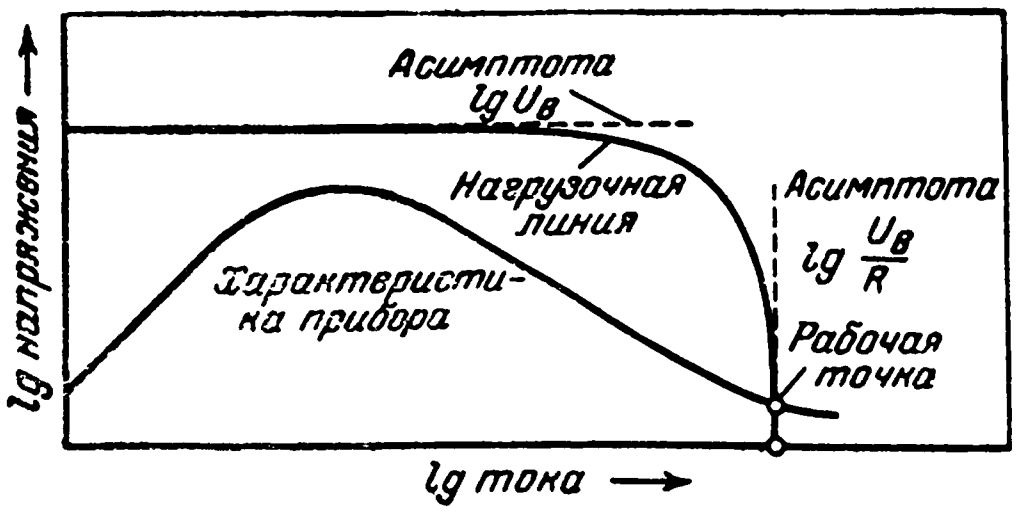


Рис. 3-19 Неожиданная форма нагрузочной линии в логарифмических координатах.

чей точке. Обратите внимание, что для этого величина положительного сопротивления R должна быть больше, чем отрицательное значение наклона характеристики элемента в рабочей точке.

В логарифмическом масштабе нагрузочная прямая преобразуется в кривую, показанную на рис. 3-19. Эта кривая имеет асимптоту напряжения U_B и асимптоту тока U_B/R.

ЗАДАЧИ

1. Покажите аналитически, что график $\lg U = f(\lg I)$ для характеристики термистора должен выражать омическое сопротивление (иметь наклон 45°) при достаточно высоком входном токе и напряжении.

2. Дано сопротивление R_0 термистора при окружающей температуре T_0, коэффициент теплоотвода элемента и величина B для материала, из которого он сделан. Разработайте метод расчета полной логарифмической характеристики термистора.

3. На основе данных, приведенных на рис. 3-3, покажите, что коэффициент теплоотвода термистора равен 0,0005 *вт/град.*

4. Для семейства временных характеристик на рис. 3-11 объясните, почему имеется «инкубационный период» между моментом замыкания цепи и началом роста тока к его конечной величине.

5. Разработан новый полупроводник, удельное сопротивление которого при комнатной температуре равно 0,1 *ом · см*; его удельное

сопротивление при 1 000° абс. равно 0,001 *ом·см* В пределах этих температур его поведение соответствует полупроводнику с собственной проводимостью.

а) Определить величину B для этого материала.

б) Из этого материала изготовлен термистор, имеющий $R_0 = 50\,000$ *ом* при комнатной температуре и $G = 0{,}0005$ *вт/град*. Интерполяцией между кривыми на рис. 3-7 определить сопротивления переменному и постоянному току этого термистора при токе подогрева 10 *ма*.

6. Все термисторы, характеристики которых приведены на рис 3-7, имеют одинаковые величины R_0, G и T_0, но отличаются по величине B, соответствующей разным отрицательным температурным коэффициентам сопротивления. Допустим, что нужно разработать материал, имеющий те же R_0, G и T_0, но большой положительный температурный коэффициент. Набросайте форму его логарифмической характеристики, которую он, по Вашему мнению, должен иметь. Видели ли Вы когда-нибудь раньше такую характеристику? Для чего, по Вашему мнению, такой материал может быть использован?

ЛИТЕРАТУРА

1. Эта глава написана по материалам: Properties and Uses of Thermistors by J. A. Becker, C. B Green, and G. L Pearson, Transactions of the American Institute of Electrical Engineers, 1946, vol. 65, p. 713; Bell System Technical Journal, 1947, vol 26, p. 170.

2. Bollman J H., Kreer J. G., The Applications of Thermistors to Control Networks, Proceedings of the Institute of Radio Engineers, 1950, vol. 38, p. 20.

ГЛАВА ЧЕТВЕРТАЯ

ПРИМЕНЕНИЕ ТЕРМИСТОРОВ

Эта глава помещена вслед за главой о термисторах для того, чтобы дать некоторое понятие о многих функциях, для выполнения которых предназначаются термисторы, и некоторое представление о важности термисторов для электронных схем. Как было указано во вводной главе, приводимые здесь описания ни в коей мере не являются полными; они дают лишь сведения, достаточные для того, чтобы читатель, знакомый с электронными схемами, мог в каждом случае представить себе применение и роль, которую термисторы играют в схемах. Помещенные здесь принципиальные схемы приведены только для иллюстрации этих применений, но отнюдь не в качестве руководящего материала для составления практических схем.

Характеристики термистор

Типы термисторов	Сопротивление, *ом*, при темпера			
	—25	0	+25	+50
Термистор A. Бусинка в стеклянной оболочке ∅0,4 *мм*	—	5 000	2 000	900
Термистор B. Бусинка в кончике стеклянного зонда	—	325 000	100 000	33 000
Термистор C. Цилиндрический стержень ∅1,25 *мм*, длиной 30 *мм* для радиозонда	87 500	37 500	18 000	9 700
Термистор D. Цилиндрический стержень ∅3,8 *мм*, длиной 18 *мм*. Общего назначения	610 000	153 000	48 400	17 300
Термистор E. Диск для непосредственного припаивания ∅5 *мм*, толщиной 2,5 *мм*	—	490	175	71
Термистор F. Диск для влажной среды ∅14,2 *мм*, толщиной 7,9 *мм*	13 000	3 200	950	340

4-1. ТЕРМОМЕТРИЯ

Наиболее очевидной областью применения термисторов, учитывая сравнительно большой температурный коэффициент сопротивления полупроводников, являются измерение и контроль температуры. Поскольку температурный коэффициент термометра с термистором может более чем в 10 раз превышать температурный коэффициент обычного металлического термометра сопротивления, термометр с термистором обеспечивает относительно большую чувствительность в аппаратуре примерно одинаковой сложности. Термисторы, подвергнутые предварительному искусственному старению, могут применяться для измерения температур до 300° С с превосходной стабильностью температурных характеристик в течение длительного времени.

Термисторный термометр, как и термометр сопротивления, обычно включается в одно из плеч моста Уитстона. Разумеется, для получения надежных результатов цепь должна быть рассчитана так, чтобы ток, проходящий через термистор, был бы достаточно мал и не

Таблица 4-1

ных термометров

турах, °С				Максимальная допустимая температура, °С	Коэффициент теплоотдачи, *мвт/град*		Тепловая постоянная времени, *сек*	
+100	+150	+200	+300		в воздухе	в воде	в воздухе	в воде
250	95	—	—	150	0,1	—	1	—
6 000	1 600	500	80	300	1,0	7	30	4
3 700	—	—	—	100	7	—	25	—
3 400	870	—	—	150	7	—	60	—
16	4,5	1,6	—	200	—	—	—	—
70	—	—	—	100	20	—	—	—

вызывал нагрева, сравнимого с наименьшим изменением окружающей температуры, которое должно быть отмечено.

Поскольку термисторные термометрические элементы могут быть изготовлены с бóльшими собственными сопротивлениями, чем металлические термометры сопротивления или термопары, сопротивления соединительных проводов обычно пренебрежимо малы сравнительно с сопротивлением элемента. Вследствие этого сам термистор может быть значительно удален от места размещения индикаторной аппаратуры. Эта возможность имеет особое значение в тех случаях, когда показания температур от значительного числа датчиков, расположенных в различных местах, должны быть сведены вместе в некотором регистрирующем или контрольном пункте.

В табл. 4-1 приведены характеристики некоторых типов термисторов, разработанных для термометрии. Термистор А бусинкового типа в стеклянной оболочке с элементом диаметром 0,4 *мм* используется для разнообразных измерений. Малые размеры делают его пригодным

для измерения температуры в весьма ограниченных пространствах. Термистор B — также бусинкового типа — расположен в кончике изолирующего стеклянного зонда длиной около 25 *мм*. Этот прибор предназначен для внутривенного измерения температуры или для измерений температуры проводящих или вызывающих коррозию жидкостей. Термистор C—стержневой цилиндрический—используется как датчик температуры в метеорологических шарах-радиозондах. Эти устройства непрерывно передают данные о температуре от шара-зонда на базовую станцию. Термистор D — стержневой цилиндрический — для различных измерений в воздухе и неактивных газах. Термистор E—небольшой диск, предназначенный для непосредственного припаивания к металлической поверхности. Он применяется для измерения температуры вентиляционных каналов, металлических резервуаров или хранилищ для жидкостей, головок цилиндров, патрубков двигателей и т. д. F представляет собой дисковый термистор, предназначенный для тропических условий, имеющий противоплесневое покрытие.

Включенный в цепь моста постоянного тока без каких-либо мер предосторожностей в отношении стабилизации температуры плеч моста, кроме обычных, с нульгальванометром чувствительностью $2 \cdot 10^{-10}$ *а·мм*, термисторный термометр легко отметит изменение температуры на 0,0005° C.

4-2. БОЛОМЕТРИЯ С ПОМОЩЬЮ ТЕРМИСТОРОВ

Если объект, температуру которого необходимо измерить, физически недоступен, находится в движении или имеет слишком высокую температуру для контактной термометрии, часто может быть использовано термисторное устройство для оптической болометрии [Л. 1]. Тепловое излучение объекта улавливается зеркалом и фокусируется на термисторе, который в таких случаях выполняется в виде листочка. Количественное определение температуры с помощью такого устройства затруднительно, за исключением случая, когда возможна градуировка для данного термистора, данной оптической системы и данного нагретого объекта, температура которого должна быть измерена. По этой причине болометрия с помощью термисторов широко применяется для обнаружения излучения, но не для измерения температуры из-

лучателя. Термисторная болометрия нашла себе практическое применение в спектроскопии [Л. 2] для обнаружения спектральных линий с помощью инфракрасных спектрометров и в военном деле для обнаружения в поле зрения нагретых объектов.

На рис. 4-1 приведена схема индикатора излучения. Излучение фокусируется на термисторе, имеющем форму листочка, поверхность которого зачернена для луч-

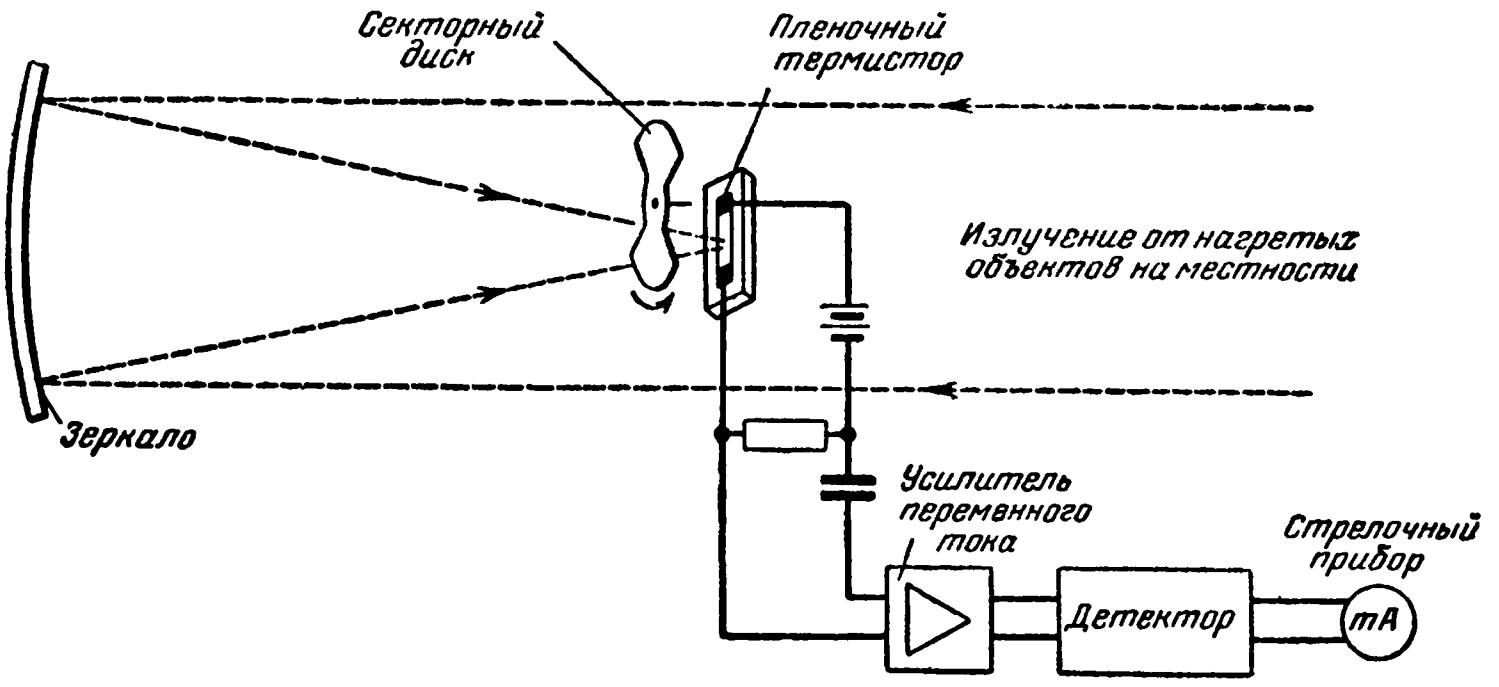


Рис. 4-1. Пленочный термистор как индикатор излучения нагретых объектов, находящихся в поле зрения.

шего поглощения излучения и нагрева листочка. Если требуется обнаружение предельно слабого излучения, цепь термистора должна быть сделана не зависящей от нестабильности напряжения источников питания, местных изменений температуры и т. д.; в таком случае применяется метод усиления с помощью переменного тока. В этом методе приходящее излучение механически прерывается перед тем, как оно падает на термисторный листочек. Это прерывание осуществляется с помощью вращающегося секторного диска, как показано на рис. 4-1. Если оптическая система направлена на нагретый объект, термистор попеременно нагревается и охлаждается и в его выходном сигнале появляется соответствующая переменная составляющая. Эта составляющая усиливается, выпрямляется и измеряется миллиамперметром постоянного тока. Таким образом, индикатор тепла дает выходной сигнал, когда он направлен на объект, температура которого отлична от местной температуры лопастей оптического прерывателя. Положение нагретых объектов

в поле зрения устанавливается затем сканированием поля зрения оптической системой индикатора; при этом в те моменты, когда выходной прибор дает показания, отличные от фона, отмечается направление луча зрения.

В течение войны были разработаны тепловые индикаторы, способные различить человека на холодном окружающем его фоне на расстоянии в несколько сотен метров или горячую дымовую трубу корабля на горизонте, если даже туман или темнота мешали прямому визуальному наблюдению.

Использование метода усиления на переменном токе заставляет делать термисторный элемент с достаточно малой постоянной времени, чтобы он успевал нагреваться и охлаждаться при прерываниях входящего излучения. Так как термистор в виде листочков имеют незначительную тепловую инерцию и сравнительно высокий коэффициент теплоотдачи в результате укрепления в хорошем тепловом контакте с теплоотводящим основанием, постоянная времени их достаточно мала для работы при обычной частоте прерывания (15—400 *гц*).

4-3. ТЕМПЕРАТУРНАЯ КОМПЕНСАЦИЯ

Естественным и очевидным применением термисторов является также компенсация изменения сопротивления электрических цепей (металлических проводников), возникающего из-за изменения температуры окружающей среды. Поскольку температурный коэффициент термисторов отрицателен, тогда как у металлических проводников он положителен, включение термистора последовательно с металлическими проводниками образует цепь, имеющую почти постоянное сопротивление в некоторых пределах температур. На рис. 4-2 изображен графически такой случай последовательной компенсации. Заметим, что сопротивление такой цепи можно считать постоянным только при отклонении температуры примерно на 10° С в обе стороны от выбранной температуры компенсации, которая в нашем случае равна 20° С.

Если, как это обычно имеет место, желательна компенсация в большем интервале температур, требуются более сложные схемы. На рис. 4-3 показана последовательно-параллельная схема компенсации. Здесь термистор включен в цепь параллельно с другим металлическим сопротивлением, величина которого примерно вдвое

меньше сопротивления компенсируемой металлической цепи. На графике приведены температурные характеристики различных элементов этой цепи, из которых видно, что сопротивление компенсированной цепи поддерживается в пределах ±2% в интервале температур от —20 до +80° С.

Сходно с описанным выше применение термистора для улучшения постоянства коэффициента самоиндукции

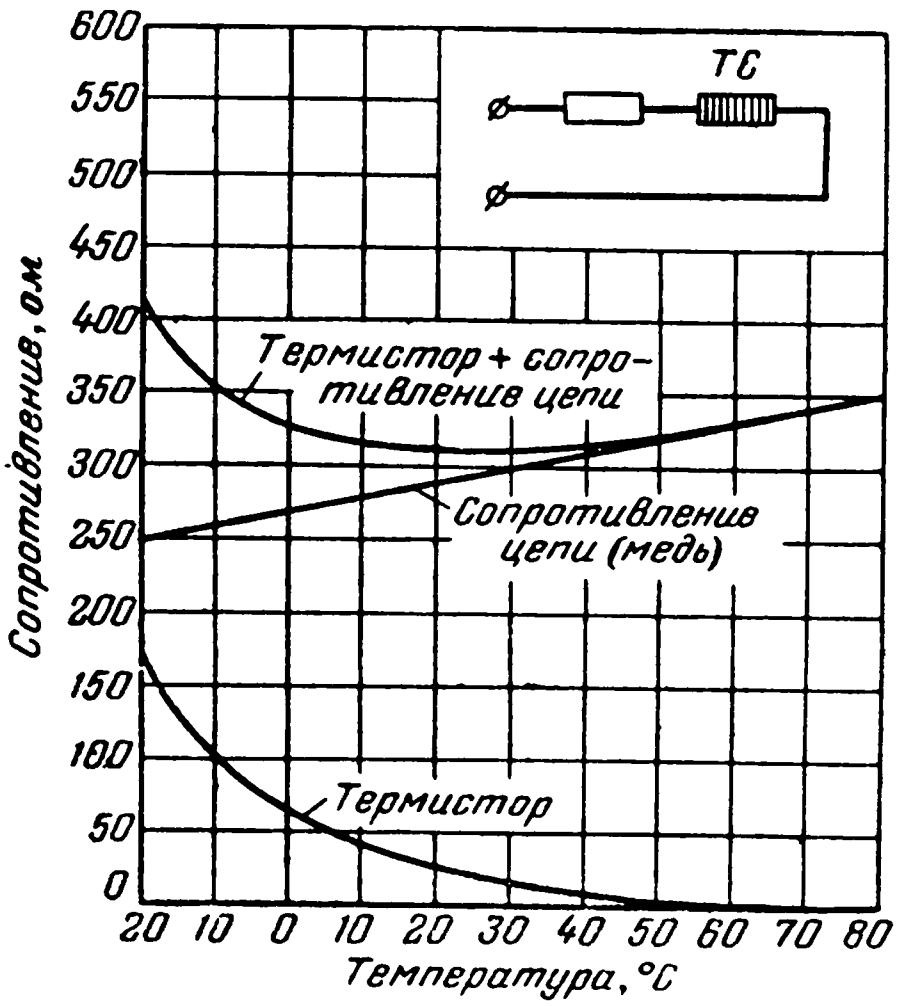


Рис. 4-2 Цепь с термистором, сопротивление которой практически постоянно в определенной области температур

катушки в пределах определенного интервала температур, даже если магнитные характеристики материала сердечника катушки изменяются в зависимости от температуры. В таких случаях сердечник частично насыщается постоянным током, пропускаемым через специальную обмотку на сердечнике. Величиной этого тока управляет термистор, заделанный в сердечник; это управление рассчитывается так, чтобы индуктивное сопротивление главной катушки приблизительно не зависело от температуры.

При разработке схем компенсации следует принять меры к тому, чтобы термистор находился в тех же температурных условиях, что и части цепи, подлежащие

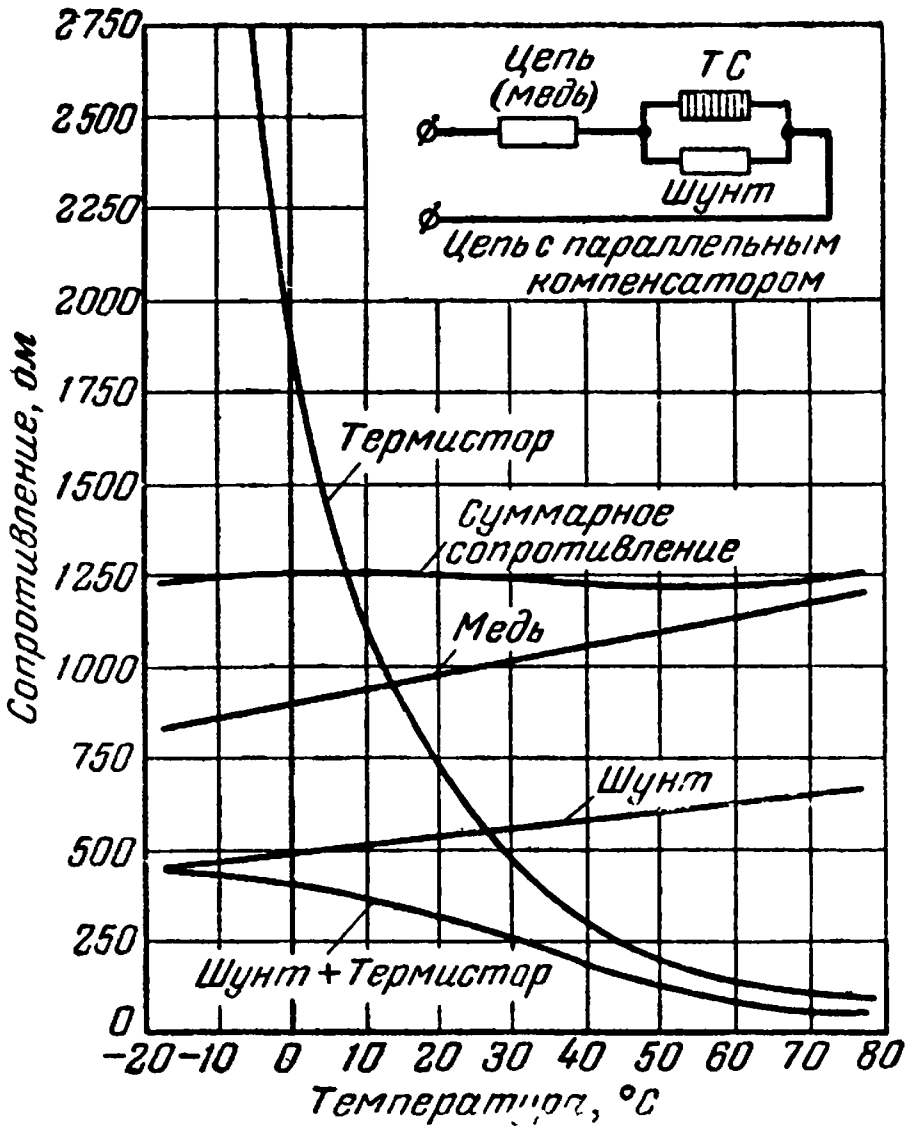


Рис. 4-3 Цепь с термистором, общее сопротивление которой постоянно в пределах значительно более широкой области температур, чем в предыдущем случае.

компенсации. Кроме того, надо учесть выделение мощности в самом термисторе и ограничить его такой величиной, когда выделяющееся джоулево тепло не вызывает значительного повышения температуры термистора, либо включить его в расчет термистора и схемы, в которой он работает.

4-4. ИЗМЕРЕНИЕ ВЫСОКОЧАСТОТНОЙ МОЩНОСТИ

Термисторы применяются в испытательных установках для измерения высокочастотной мощности [Л. 3]. Чувствительность таких установок на порядок выше чувствительности металлических высокочастотных термопар и их легче изготовить с собственными сопротивлениями, согласующимися с полными сопротивлениями волноводов и высокочастотных линий передач.

В качестве измерителя мощности термистор включается в конец линии или волновода, подводящего подлежащую измерению мощность. Оконечное устройство рассчитывается так, что термистор поглощает всю или из-

вестную часть подводимой мощности. Выделяющееся при этом джоулево тепло нагревает термистор и, таким образом, понижает его сопротивление.

На рис. 4-4 показан термистор в типичной схеме для измерения мощности. Термистор является одной ветвью моста Уитстона и находится под таким постоянным смещением, что рабочая точка располагается справа от максимума его характеристики, где его сопротивление легко может быть согласовано с нагрузкой волновода. Мост

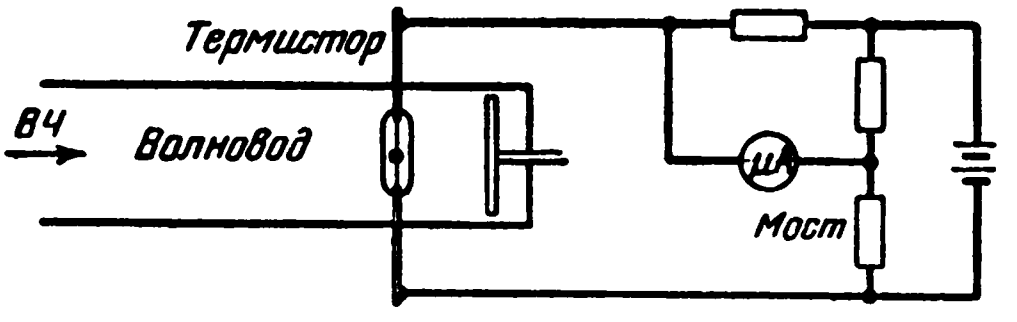


Рис. 4-4. Термистор в схеме измерения мощности, поступающей в волновод.

балансируется на нуль в отсутствие высокочастотной мощности, подача которой вызывает разбаланс моста, зависящий от ее величины.

В типичной испытательной установке такого типа отклонение индикатора моста на всю шкалу соответствует высокочастотной мощности 1 *мвт*. Индикатор представляет собой панельный измерительный прибор.

Если такие измерители мощности рассчитываются для работы в широких пределах изменений температуры окружающей среды, то схема должна быть температурно компенсирована введением еще двух термисторов, не показанных на рис. 4-4. Эти добавочные компенсирующие термисторы рассчитываются так, чтобы они не реагировали на ток моста, но были чувствительны к изменениям температуры окружающей среды. Один из этих компенсаторов поддерживает баланс моста, другой же сохраняет калибровку моста независимо от величины температуры.

4-5. УПРАВЛЕНИЕ ХАРАКТЕРИСТИКАМИ ЭЛЕКТРОННЫХ ЦЕПЕЙ

В системах уплотнения проводных линий дальней связи, работающих с промежуточными усилителями, качество работы всей системы сильно зависит от постоянства характеристик каждого промежуточного усилителя. Небольшое изменение характеристики передачи каждого

усилительного участка, происходящее, например, за счет изменения температуры в пределах значительной географической зоны, повторенное на большом числе включенных последовательно участков, может привести к недопустимо высокому общему изменению характеристики всей системы. Поэтому необходимо тщательное регулирование каждого участка. Применение термисторов для автоматической стабилизации характеристик систем связи имеет сравнительно давнюю, но чрезвычайно интересную историю [Л. 4, 5 и 6].

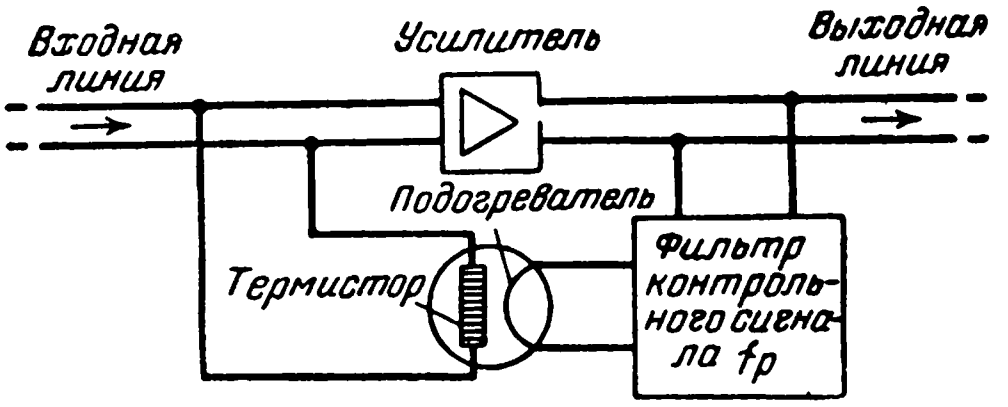


Рис. 4-5. Термистор косвенного подогрева как элемент отрицательной обратной связи в схеме стабилизации уровня выходного сигнала усилителя.

Такая регулировка может производиться изменением коэффициента усиления каждого промежуточного усилителя. В одном из вариантов таких схем термистор с косвенным подогревом присоединен к усилителю в качестве элемента обратной связи, как показано на рис. 4-5. Катушка подогрева отбирает некоторую часть выходной мощности усилителя, тогда как подогреваемый термисторный элемент шунтирует вход усилителя, вызывая некоторое затухание. Если амплитуда входного сигнала возрастает, выходная мощность увеличивается, термисторный элемент нагревается, его сопротивление падает и повышается часть входной мощности, поглощаемая термистором, обеспечивая этим подачу напряжения отрицательной обратной связи на вход усилителя. Таким образом, любое температурное изменение характеристики потерь в предшествующем усилительном участке линии в значительной степени нейтрализуется соответствующим изменением в противоположном направлении усиления усилителя. В таких системах передачи регулировка зависит не от уровня несущих частот, на которых производится передача информации, а в основном от уровня

контрольного сигнала, вводимого для этой цели на вход линии. Частота контрольного сигнала выделяется фильтром, введенным в цепь обратной связи для того, чтобы к подогревателю термистора приходил только контрольный сигнал. Тепловая инерция термистора и его подогревателя, конечно, слишком велика, чтобы термистор мог реагировать на частоту контрольного сигнала; он реагирует лишь на его огибающую.

В другом промежуточном усилителе цепь обратной связи содержит термистор прямого подогрева, включенный так, чтобы регулировать амплитуду напряжения обратной связи вместо величины затухания во входной цепи.

Такая управляемая термистором обратная связь может являться также устройством для защиты от перегрузок. Однако в таких схемах цепь обратной связи с термистором должна быть чувствительна ко всем частотам, по которым желательна защита от перегрузок. В подобной же схеме термистор может быть использован для автоматической регулировки усиления усилителя низкой частоты. В этих случаях должно быть принято во внимание, что из-за тепловой инерции термистора регулировка может быть не мгновенной, а слегка запаздывающей.

Термисторы применялись для стабилизации как амплитуды [Л. 7 и 8], так и частоты [Л. 9] генераторов колебаний низкой и высокой частот. В одном методе стабилизации амплитуды положительная обратная связь с анода на сетку генераторной лампы управляется степенью разбаланса моста, в одну из ветвей которого включен термистор. Этот термистор нагревается частью выходной мощности генератора. Чем выше температура, до которой нагрет термистор, тем ближе мост к равновесию и тем ниже напряжение возбуждения на сетке. Поэтому амплитуда выходного напряжения стабилизируется на таком уровне, при котором разбаланс моста с термистором создает на сетке сигнал возбуждения, как раз достаточный для поддержания постоянного выходного напряжения. Этот метод стабилизации более надежен, чем основанный только на нелинейности самой лампы. При этом лучше обеспечивается сохранение гармонического состава выходного сигнала и его амплитуда менее зависит от колебаний напряжения питания, изменения параметров ламп и их старения при работе.

4-6. СТАБИЛИЗАЦИЯ НАПРЯЖЕНИЯ

Термистор может быть применен в очень простой цепи в качестве стабилизатора напряжения. На рис. 4-6 показаны такая цепь и графическое построение, иллюстрирующее принцип стабилизации. Приведены характеристики термистора и последовательно с ним включенного сопротивления R_s, образующих параллельную ветвь стабилизатора. R_s представляет собой сопротивление, величина

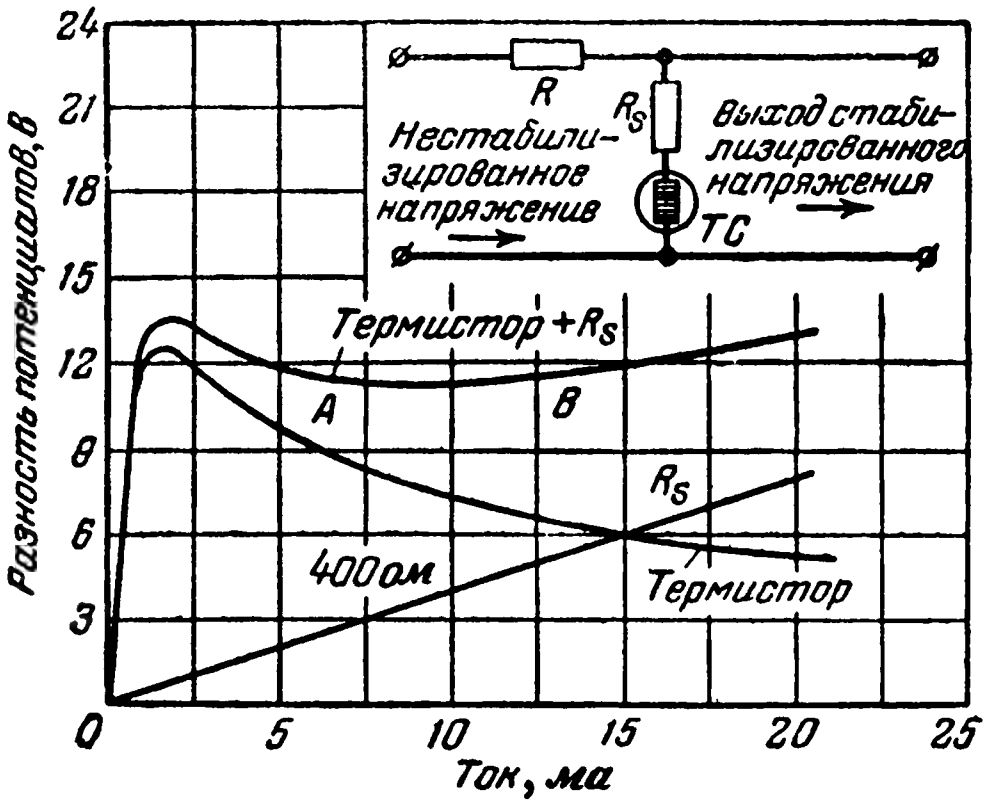


Рис. 4-6. Стабилизатор напряжения, использующий нелинейную часть характеристики термистора.

которого равна абсолютной величине отрицательного сопротивления термистора в правой от максимума части его характеристики. Характеристика цепи термистора вместе с последовательно включенным сопротивлением, представленная кривой «Термистор+R_s», имеет особую область между точками A и B, в пределах которой напряжение почти не зависит от тока. Этот участок является рабочей областью стабилизатора. Изменение входного напряжения вызывает изменение тока стабилизатора, причем колебания напряжения в основном приходятся на колебания падения напряжения IR на сопротивлении R. Подобный стабилизатор имеет то достоинство, что он поддерживает выходное напряжение при изменении не только входного напряжения, но также и выходной нагрузки и потребления тока в сравнительно широких пределах.

4-7. ИЗМЕРЕНИЕ ДАВЛЕНИЯ

Применение термисторного элемента в качестве вакуумного манометра основано на том факте, что коэффициент теплоотдачи термисторного элемента, помещенного в воздухе, понижается с уменьшением давления. На рис. 3-5 приведены кривые, из которых следует, что положение логарифмических характеристик термистора зави-

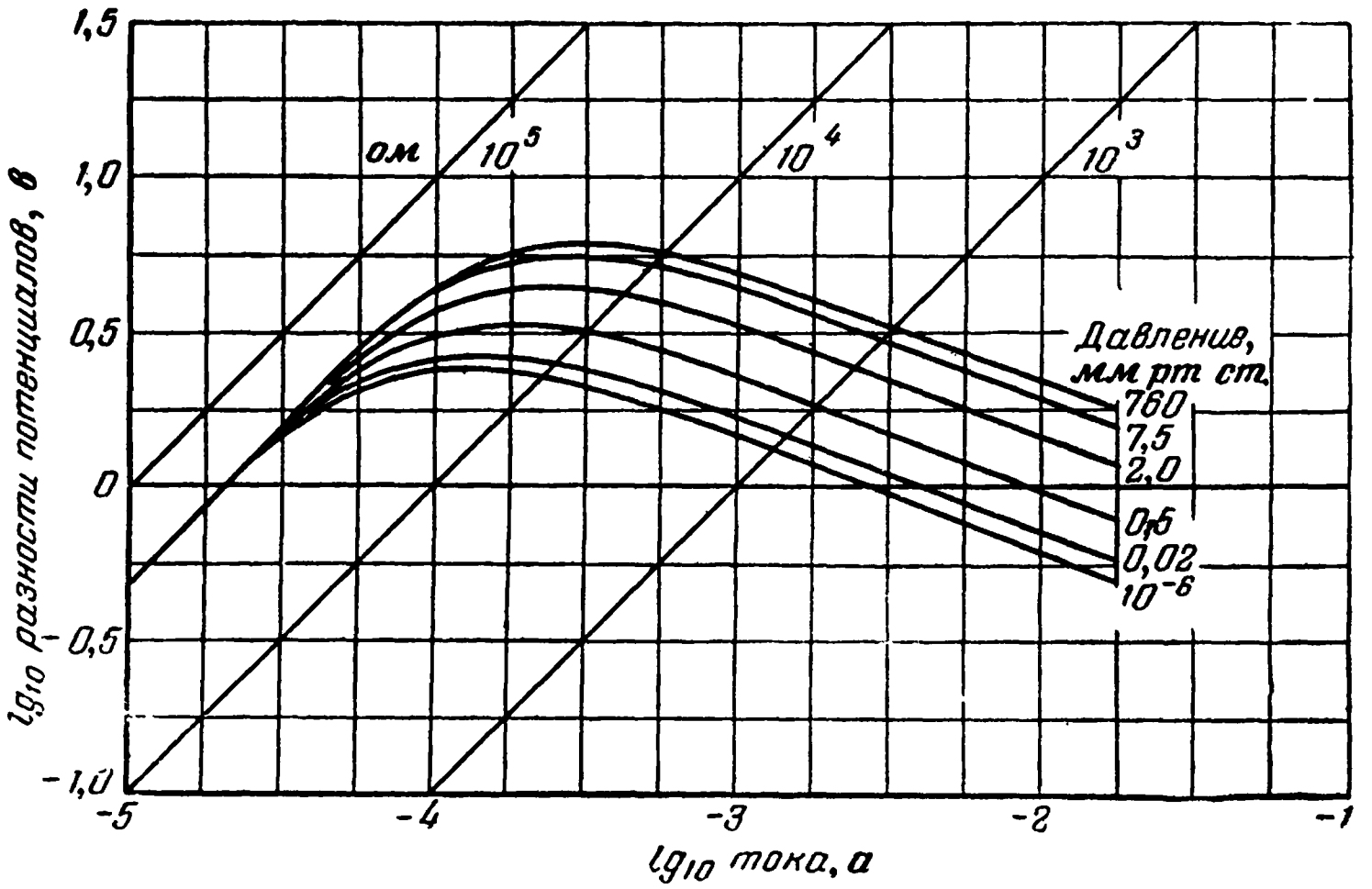


Рис. 4-7. При изменении давления окружающего газа меняется коэффициент теплоотдачи термистора, что позволяет использовать его в качестве вакуумного манометра.

сит от величины коэффициента теплоотдачи. На рис. 4-7 показано подобное же семейство кривых для бусинкового термистора диаметром 0,5 *мм* в стеклянной оболочке в спокойном воздухе при различных давлениях от атмосферного до высокого вакуума. Такое семейство кривых является основой для применения термистора как манометра. Определение давления воздуха, окружающего термистор, может быть произведено с помощью любой схемы, обеспечивающей измерение сопротивления термистора при некотором известном токе, например 1 *ма*, как на рис. 4-7.

Точность такого определения зависит от того, насколько температура манометра при работе поддерживается близкой к температуре, при которой снято семейст-

во кривых на рис. 4-7. Чтобы уменьшить влияние температуры на измерение давления термисторным манометром, обычно применяют мостовую схему, подобную показанной на рис. 4-8, со вторым идентичным термистором в одной из ветвей моста, запаянным в стеклянный баллон с высоким вакуумом. Мостик настраивается на точный баланс, когда первый термистор также находится в высоком вакууме. Если теперь первый тер-

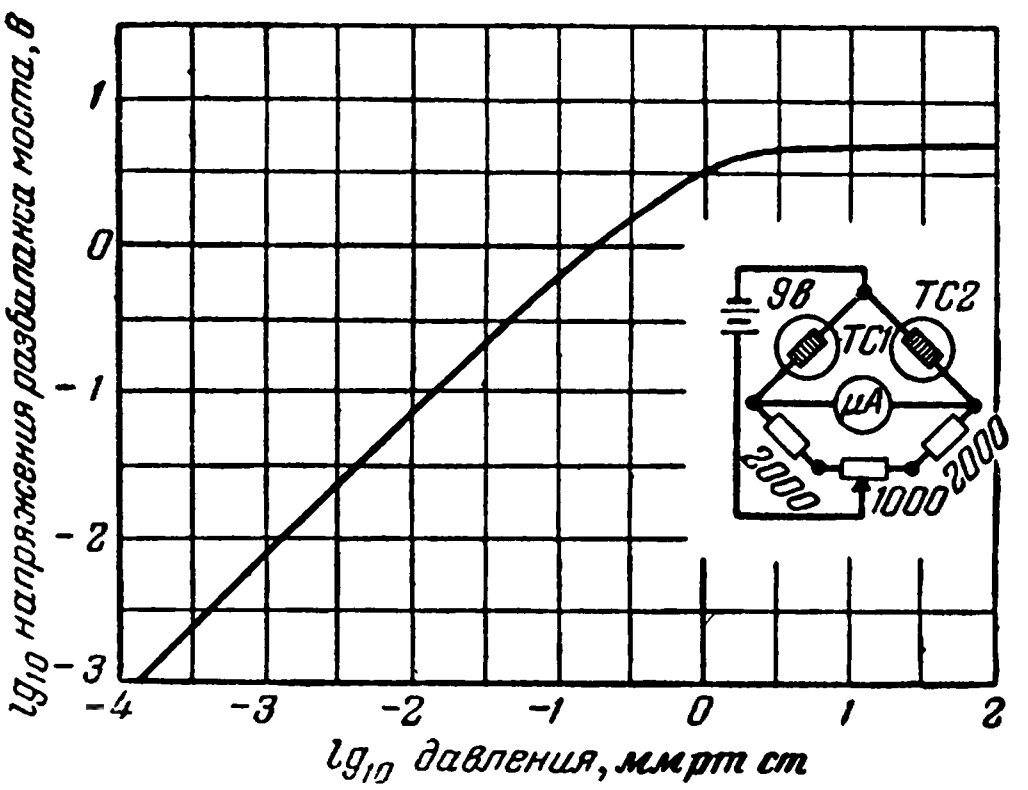


Рис. 4-8 Градуировочная кривая мостового манометра. Термистор TC_1 постоянно находится в отпаянном баллоне с высоким вакуумом; термистор TC_2 помещают в подлежащий измерению вакуум.

мистор находится под давлением, величину которого нужно измерить, то разбаланс моста даст величину давления. На рис. 4-8 представлена градуировочная кривая для типичного термисторного манометра. При настройке моста напряжение питания и отношение сопротивлений ветвей должны быть, конечно, выбраны так, чтобы рабочая точка термистора была расположена правее вершины его характеристики, где изменения коэффициента теплоотдачи вызывают значительные изменения сопротивления. Кроме того, оба термистора должны быть расположены как можно ближе один к другому, чтобы находиться в примерно одинаковых температурных условиях.

Термисторный манометр особенно ценен для тех диапазонов давлений, где прочие измерители давлений нечувствительны. Так, например, жидкостный маномегр с U-образной трубкой практически бесполезен при давле-

ниях ниже 0,1—0,2 *мм рт. ст.*, тогда как ионнизационный манометр нельзя применять при давлении выше примерно 10^{-3} *мм рт. ст.* из опасения повреждения катода ионной бомбардировкой. В принципе термисторный манометр мостового типа, подобный описываемому, должен годиться для измерения давлений в области высокого вакуума (10^{-5} *мм рт. ст.* и ниже), однако практические трудности поддержания баланса моста (установки нуля) в течение достаточно длительного времени и трудность обеспечения точного равенства температуры обоих термисторов и металлических элементов цепи ставят практический нижний предел пригодности их для измерений давлений примерно 10^{-4} *мм рт. ст.*

Близко примыкающей областью применения термисторов является измерение скоростей потоков газов и жидкостей. Коэффициент теплоотдачи термистора, погруженного в движущийся газ или жидкость, зависит от скорости потока; он выше для больших скоростей. Различие в величинах коэффициента теплоотдачи проявляется в смещении характеристик термистора; что позволяет легко отградуировать данный термистор для различных скоростей потока в различных средах.

4-8. ВЫДЕРЖКА ВРЕМЕНИ И СХЕМЫ ЗАДЕРЖКИ

Значительное применение термисторы нашли в схемах выдержки времени и задержки. Мы уже видели в § 3-3-1, что это применение основано на использовании тепловой инерции материала термисторного элемента и частей оболочки термистора, находящихся в тесном тепловом контакте с элементом. Из рис. 3-11 видно, что существует задержка во времени между моментом подачи напряжения в цепь, содержащую термистор, и достижением установившегося значения тока Так, например, если струнный гальванометр в цепи, показанной на рис. 3-11, заменить на реле и отрегулировать его на срабатывание, скажем, при 10 *ма*, то при напряжении питания 60 *в* реле сработает через 1 *сек* после замыкания ключа. При соответствующей регулировке реле и выборе напряжения питания и типа термистора можно получить задержки длительностью от нескольких миллисекунд до 10 *мин* («силовая» задержка времени). Достоинство термисторов для схем задержки этого типа заключается в их малых размерах, малом весе, надежности и большой долговечности.

Применение термисторов для «силовой» задержки наиболее целесообразно в случаях, не требующих высокой точности длительности времени задержки. В некоторых случаях допустима разница в 6 раз между самой длительной и самой короткой задержками. Такой разброс необходим для того, чтобы перекрыть складывающиеся одновременно наименее выгодные допустимые отклонения от номинальных значений следующих параметров, влияющих на длительность времени задержки: рабочего напряжения, сопротивления проводников цепи, сопротивления обмотки реле, тока срабатывания реле, температуры окружающей среды и разброса характеристик между различными экземплярами термисторов одного типа.

После осуществления задержки требуется время для охлаждения термистора перед следующим срабатыванием цепи задержки. Если такой интервал не будет предусмотрен, то вторая задержка окажется короче первой. Необходимое время охлаждения зависит от многих параметров; обычно оно в несколько раз больше длительности задержки. В релейных цепях, в которых новая задержка может потребоваться вскоре после предыдущей, термистор обычно при срабатывании реле замыкают накоротко добавочной контактной парой, имеющейся в реле. Эта мера позволяет термистору охладиться во время включения реле, в результате чего после отпускания реле термистор быстрее оказывается готовым к работе.

Чтобы избежать больших отклонений в длительности времени задержки и длительных интервалов охлаждения, более или менее неизбежных в «силовых» схемах задержки, были разработаны другие схемы цепей задержки, использующие термисторы. В одной из них задержка определяется временем, необходимым для охлаждения термистора от начальной высокой температуры до некоторой более низкой температуры, при которой происходит срабатывание реле. Этот метод получения задержки позволяет избежать «инкубационного» периода, очевидного из кривых на рис. 3-11, и, кроме того, дает возможность быстро подготовить термистор для новой задержки сразу по окончании предыдущей путем быстрого нагрева термистора до высокой температуры.

Помимо применения термисторов в качестве элемен-

тов задержки, их используют и в ряде других сходных случаев. Так, можно защитить реле от ложных срабатываний под влиянием случайных всплесков напряжения, включив термистор последовательно с обмоткой защищаемого реле. В случае такого всплеска напряжения начальное сопротивление холодного термистора ограничивает ток величиной, меньшей тока срабатывания реле, если только длительность всплеска не очень велика. С другой стороны, нормальное рабочее напряжение, действующее дольше постоянной времени тепловой инерции термистора, вызывает срабатывание реле после короткой задержки. Термистор, таким образом, позволяет отличать полезный сигнал, имеющий значительную длительность, от ложного сигнала короткой длительности, даже если амплитуда последнего в несколько раз превышает полезный сигнал.

4-9. СИСТЕМЫ ПОЖАРНОЙ СИГНАЛИЗАЦИИ

Особенности характеристик термистора для различных внешних температур (см. § 3-3) делают его идеальным элементом электрической системы пожарной сигнализации, срабатывающей от температурного воздейст-

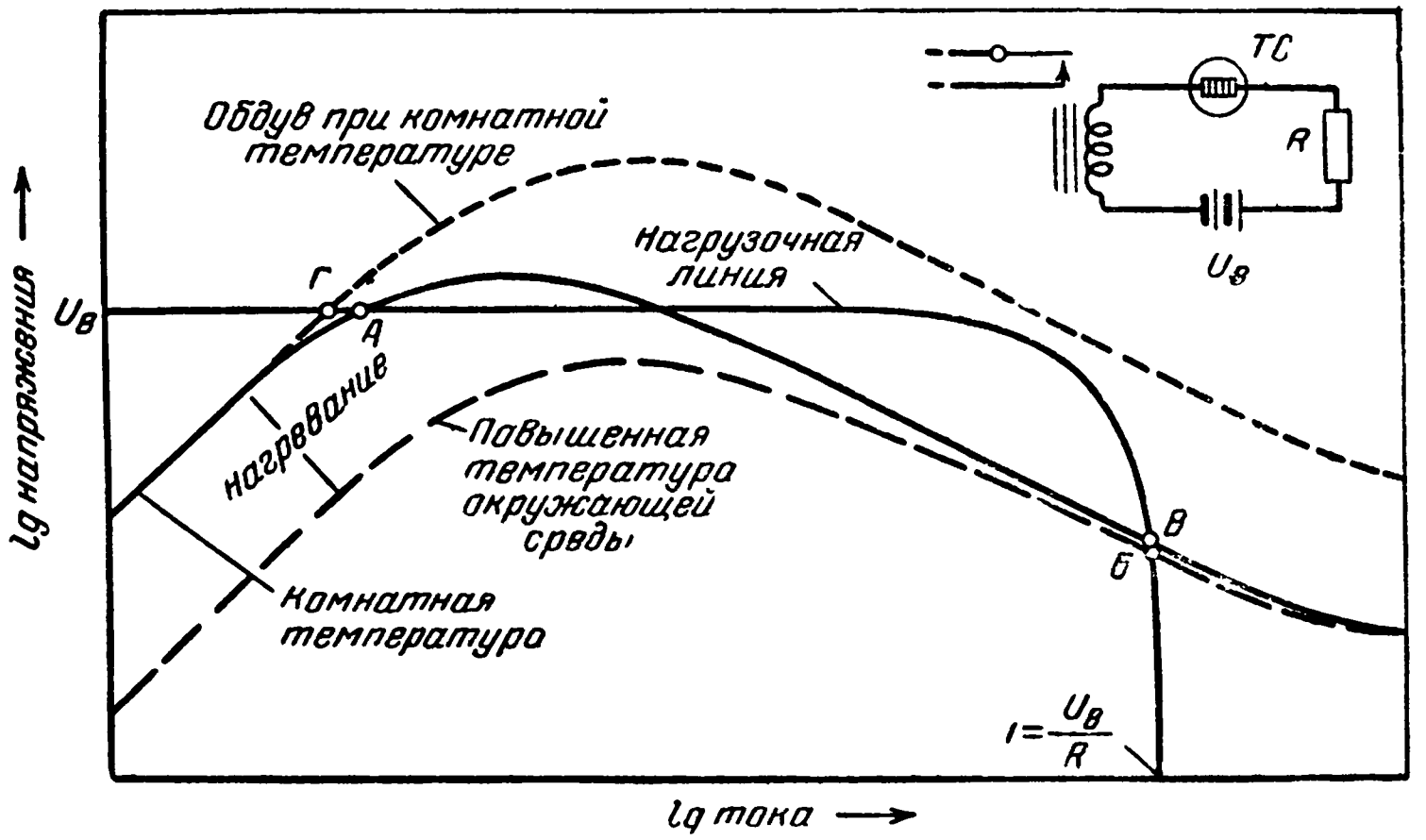


Рис. 4-9. В цепи пожарной сигнализации сигнальное реле срабатывает при скачкообразном перемещении рабочей точки из первоначального положения *А* в положение *Б*, происходящем в результате нагревания термистора.

вия. Если включить термистор последовательно с источником питания и сопротивлением нагрузки, как показано на рис. 4-9, и если напряжение питания меньше напряжения в максимуме характеристики термистора при нормальной внешней температуре, то в цепи установится режим, соответствущий рабочей точке *А* на сплошной кривой [1]. Если возникает пожар и термистор оказывается в атмосфере, более теплой, чем раньше, то максимум характеристики термистора опускается ниже горизонтальной части нагрузочной линии. Рабочая точка цепи вследствие этого перескакивает вправо в точку *Б* (на изображенной длинными штрихами характеристике), что может соответствовать току цепи, в несколько сотен раз большему, чем в точке *А*. Увеличившийся ток вызывает срабатывание реле, включающего сигнал тревоги.

Если пожар погашен и температура окружающего пространства возвращается к первоначальной нормальной величине, ток в цепи сигнализации меняется лишь незначительно за счет сдвига рабочей точки из *Б* в *В* на рис. 4-9. Рабочая точка может быть возвращена в начальное положение, соответствующее малому току в точке *А*, либо посредством кратковременного размыкания цепи питания, либо обдувом термистора.

Обдув воздухом, быстро обтекающим термистор, значительно увеличивает его коэффициент теплоотдачи и смещает характеристику в положение, указанное короткими штрихами. Если это смещение имеет достаточную величину, то характеристика термистора может уйти за пределы изогнутой части нагрузочной линии и тогда останется единственная рабочая точка *F*. После прекращения обдува цепь приходит в свое первоначальное состояние, соответствующее точке *А*. Такую цепь можно назвать тепловым триггером. Действительно, эта цепь может находиться в двух устойчивых состояниях по току и может быть переведена скачком из одного из этих состояний в другое с помощью тепловых воздействий.

---

[1] Нагрузочная линия на рис. 4-9 может поразить своей необычной формой тех, кто привык при графическом анализе пользоваться обычными характеристиками приборов и прямыми нагрузочными линиями Однако мы сейчас пользуемся двойными логарифмическими координатами, в которых нагрузочная линия правильно изображается так, как показано на этом графике (см § 3-6)

4-10. УСИЛИТЕЛИ С ОТРИЦАТЕЛЬНЫМ СОПРОТИВЛЕНИЕМ

Отрицательное сопротивление переменному току, характерное для термистора при положении рабочей точки справа от максимума напряжения на его характеристике, может быть использовано в усилителе с отрицательным сопротивлением. На рис. 4-10 показана схема такого усилителя; отрицательное сопротивление в нем используется для частичной компенсации положительного сопротивления ветви, содержащей генератор и нагрузку.

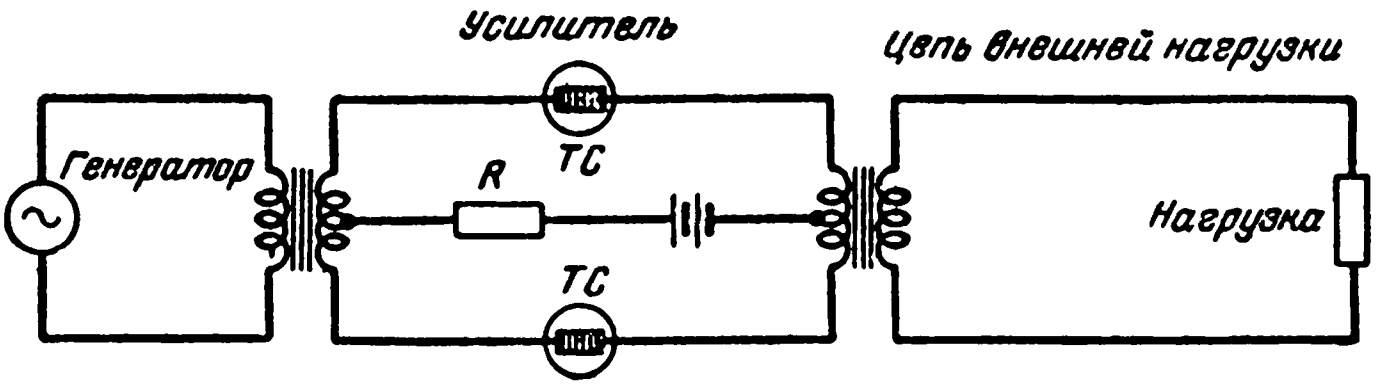


Рис 4-10. Два термистора в этой цепи действуют как сбалансированный усилитель с отрицательным сопротивлением

Усилитель имеет два одинаковых термистора в цепи источника питания, включенных последовательно с сопротивлением, стабилизирующим положение рабочих точек термисторов при значениях постоянного напряжения и тока, соответствующих части характеристики с отрицательным наклоном. При этом для переменных сигналов, проходящих от генератора к нагрузке, термисторы являются отрицательными сопротивлениями. Усилитель получается сбалансированным, поскольку ветвь цепи, содержащая источник питания термисторов, включена между средними точками входного и выходного трансформаторов, вследствие чего переменная составляющая сигнала минует источник питания и стабилизирующее сопротивление, в то время как изоляция трансформаторов не допускает гальванической связи с остальной частью цепи.

Для устойчивости всей схемы необходимо, чтобы отрицательное сопротивление усилителя не превышало положительного сопротивления остальной части схемы, иначе система войдет в режим самовозбуждения на частоте, определяемой реактивными сопротивлениями цепей. Необходимость обеспечения достаточного запаса устойчивости ограничивает усиление, надежно обеспечиваемое таким усилителем с отрицательным сопротивлением ве-

личиной примерно 10—12 *дб*. Конечно, полоса частот, в пределах которой может происходить усиление, ограничена сверху критической частотой термистора. При желании иметь достаточно плоскую частотную характеристику усилителя в пределах рабочей полосы частот ее верхнюю частоту надо выбирать по крайней мере в 10 раз меньшей критической частоты термистора.

ЗАДАЧИ

1. Термистор, характеристики которого показаны на рис. 3-3, включен последовательно с батареей напряжением 24 *в* и сопротивлением 2 400 *ом*. а) Каков ток в цепи в момент включения? б) Какова величина установившегося тока после того, как термистор полностью прогреется? в) Какова установившаяся температура термистора? г) Каковы сопротивления термистора постоянному и переменному току в окончательно установившемся режиме?

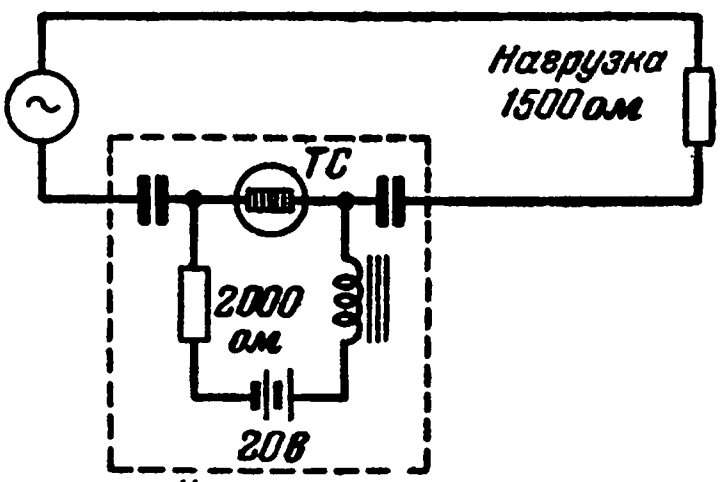


Рис 4-11. Схема к задаче 3

2. Термистор, сделанный из материала № 1 ($B=4\,000$) и имеющий $R_0=100\,000$ *ом* при комнатной температуре, используется как термометр сопротивления. Любым методом определите тепловую чувствительность $\Delta R/\Delta T$ термистора при комнатной температуре.

3. Рассчитайте усиление термисторного усилителя с отрицательным сопротивлением, схема которого помещена на рис. 4-11. Характеристики термистора приведены на рис. 3-3. Примите равными нулю: выходное сопротивление переменному току генератора сигнала, емкостное сопротивление конденсаторов, сопротивление постоянному току индуктивностей; индуктивное сопротивление последних примите бесконечным. Частота сигнала намного ниже критической частоты термистора.

4. Изготовлен стабилизатор напряжения по схеме рис. 4-6, где $R=2\,000$ *ом* и $R_в=750$ *ом*, с термистором, характеристики которого показаны на рис. 3-3 Как будет изменяться выходное напряжение ненагруженного стабилизатора при изменении напряжения питания до величин: 20 *в*; 22,5 *в*; 25 *в*; 27,5 *в*; 30 *в*? Повторите расчет для выходной нагрузки 5 000 *ом*

5. Термистор, характеристики которого приведены на рис. 3-3, включен последовательно с сопротивлением 1 000 *ом*. Определите характеристику этой цепи для напряжений питания от нуля до 20 *в* как при возрастании, так и при убывании напряжения.

ЛИТЕРАТУРА

1. Brattain W. H., Becker J. A., Thermistor Bolometers, Journal of the Optical Society of America, 1946, vol. 36, p. 354.

2. Kivenson G., Steinback R. T., Rider M., An Infra-

Red Chopped Radiation Analyzer, Journal of the Optical Society of America, 1948, vol. 38, p. 1086.

3. Gaffney F. J., Microwave Measurements and Test Equipments, Proceedings of the Institute of Radio Engineers, 1946, vol. 34, p. 775.

4. Bollman J. H., A Pilot-Channel Regulator for the K-1 Carrier System, Bell Laboratories Record, 1942, vol. 20, p. 258.

5. Tweeddale J. E., Thermistors, Western Electric Oscillator, December 1945, vol. 3, p 34.

6. Johnson J. C., Thermistor Techniques, Electronic Industries, 1945, vol. 4, p. 74.

7. Meacham L., The Bridge-Stabilized Oscillator, Proceedings of the Institute of Radio Engineers, 1938, vol. 26, p. 1278.

8. Shepherd W. G, Wise R O., Frequency-Stabilized Oscillators, Proceedings of the Institute of Radio Engineers, 1943, vol. 31, p. 256.

9. Guené P, Utilisation de thermistances pour compenser la dérive thermique de montages auto-oscillateurs, Annales de radioélectricité, 1956, vol 11, p. 317.

ГЛАВА ПЯТАЯ

ВАРИСТОРЫ

ВВЕДЕНИЕ И ОПИСАНИЕ ПЛОСКОСТНЫХ ПРИБОРОВ[1]

Варисторами, или «переменными сопротивлениями» (variable resistors), называют приборы, сопротивление которых зависит от напряжения, тока или полярности. Мы уже встречались с нелинейностью характеристик термисторов; механизм этой нелинейности носил тепловой характер. У варисторов, напротив, причина их электрической нелинейности не имеет ничего общего с тепловыми явлениями, и физика варисторов представляет собой самостоятельную область изучения.

На долю варисторов приходится львиная доля продукции полупроводниковых устройств. Точечные диоды применяются как детекторы модулированных высокочастотных сигналов, амплитудные селекторы, ограничители амплитуды, выпрямители для токов порядка микроампер и миллиампер, элементы ключевых схем, управляемые постоянным током, устройства для «привязки» уровня сигнала или восстановления постоянной составляю-

---

[1] Термин «плоскостной» использован для того, чтобы отличать эти приборы от точечных, которые рассмотрены в следующей главе.

щей и в качестве логических схем. Диоды с большой контактной площадью используются как выпрямители с токами в пределах от десятков миллиампер до десятков ампер. Варисторы широко используются как модуляторы и демодуляторы в системах уплотнения проводных линий дальней связи, как экспандеры и компрессоры динамического диапазона сигналов и для защиты от перегрузок при малых уровнях напряжения. Кремниевые диоды сплавного типа применяют в тех случаях, когда требуется очень большое отношение прямого сопротивления к обратному. Симметричные варисторы применяются там, где нужна нелинейная характеристика в диапазоне напряжений от 1 *в* до десятков киловольт (область, используемая в большинстве случаев защиты от перегрузок).

5-1. НЕСИММЕТРИЧНЫЕ ВАРИСТОРЫ

Твердый выпрямитель (или диод) представляет собой варистор, характеристика которого не только нелинейна, но и несимметрична относительно начала координат. Форма характеристики выпрямителя показывает, что прибор легко пропускает ток в одном направлении, но представляет большое сопротивление току в другом направлении. Эти приборы объединяются под своеобразным названием несимметричных варисторов.

Полупроводниковый контактный выпрямитель независимо от его размера, формы и электрических характеристик состоит из слоя или куска полупроводникового материала с двумя металлическими контактами на его поверхности (или поверхностях). Если бы оба эти контакта обладали выпрямительными свойствами, то такой прибор представлял бы собой просто два выпрямителя, включенные последовательно в противоположных направлениях. Поэтому, чтобы лучше использовать свойства прибора, стараются возможно более приблизить выпрямительные свойства одного из этих контактов к идеальным, в то же время подавляя или устраняя выпрямляющее действие другого контакта. Это может быть осуществлено выбором для второго контакта подходящего металла, контакт которого с данным полупроводником не обладает выпрямительными свойствами, или обработкой поверхности полупроводника перед наложением металлическского контакта, уничтожающей

выпрямительное действие. На рис. 5-1 показана в разрезе идеализированная картина полупроводникового выпрямителя, изготовленного подобным образом.

Выпрямляющая область нелинейного контакта часто называется барьерным или запирающим слоем. Она состоит из непосредственно прилегающего к металлическому контакту тонкого слоя полупроводника, местное удельное сопротивление которого во много раз больше, чем во всей остальной массе, а величина этого удельного сопротивления, так же как и толщина самого слоя, зависит от величины и направления приложенного к кон-

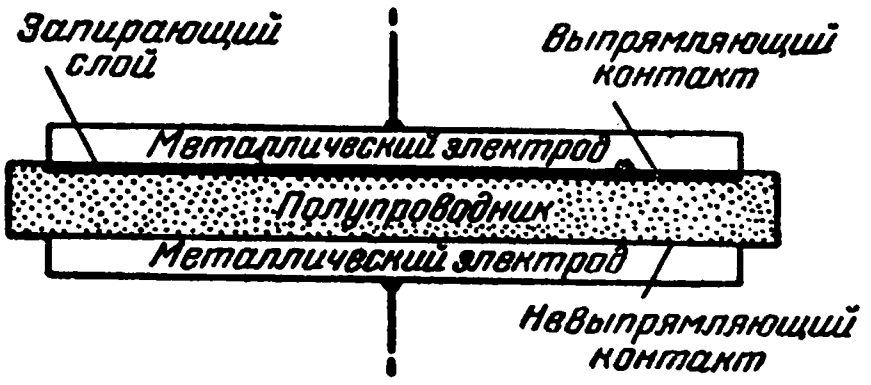


Рис 5-1. Схематический разрез полупроводникового контактного выпрямителя, состоящего из слоя полупроводника между выпрямляющим и невыпрямляющим контактами с металлическими электродами

тактам напряжения. Толщина типичных запирающих слоев лежит в пределах 10^{-6}—10^{-3} *см*. Теория запирающего слоя изложена в гл. 19 и 20.

Основная функция выпрямителя — однонаправленный вентиль, пропускающий ток свободно в одном направлении и плохо в другом.

Идеальный выпрямитель должен был бы, конечно, иметь бесконечную проводимость в одном направлении и бесконечное сопротивление в другом. Выпрямительные устройства различных видов, основанные на использовании контактов между металлами и полупроводниками, приближаются к этому идеалу с некоторыми характерными отличиями, более или менее общими для всех приборов такого рода. Эти общие особенности поясняет рис 5-2, на котором показана характеристика меднозакисного варистора. Характеристика эта весьма нелинейна и чрезвычайно несимметрична. В направлении свободного прохождения тока, называемом прямым на-

правлением, ток возрастает быстрее, чем пропорционально напряжению.

У некоторых купроксных выпрямителей характеристика для прямого направления имеет ограниченный участок, где ток возрастает пропорционально четвертой степени напряжения. Именно на этом участке, где нелинейность наибольшая, располагают рабочую точку выпрямителя, когда он служит детектором слабых сигналов

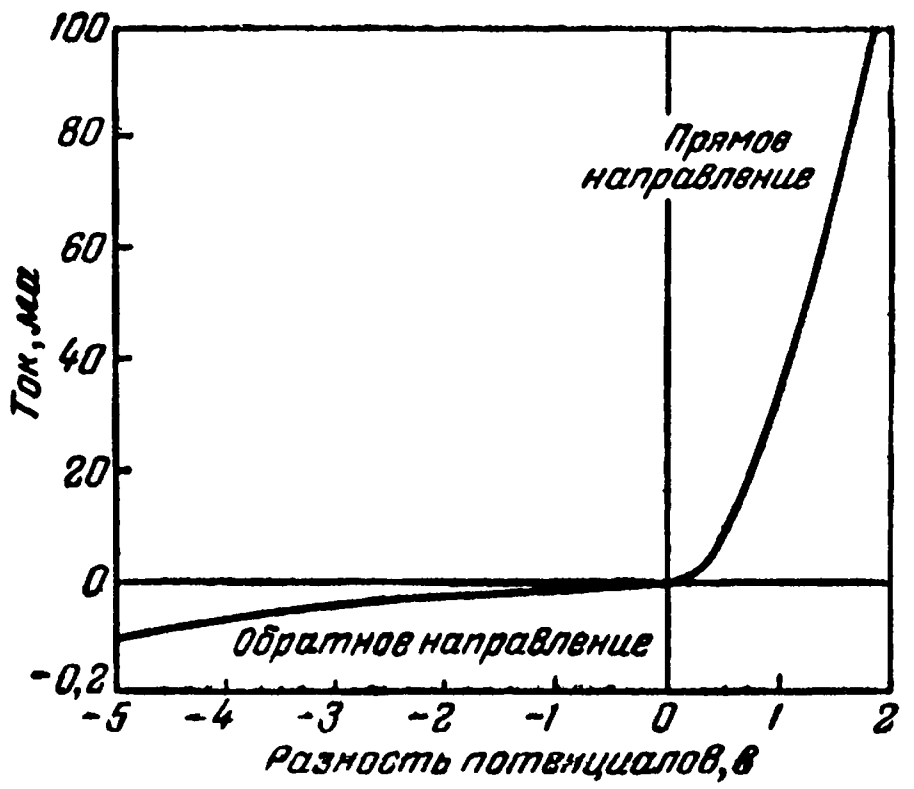


Рис. 5-2. Характеристика дискового медно-закисного выпрямителя диаметром ~ 5 *мм*
Для прямого и обратного токов масштабы на осях координат разные

переменного тока. В другом направлении, называемом обратным, при всех напряжениях протекает слабый ток утечки (обратный ток). Если обратное напряжение превысит в случае меднозакисного выпрямителя примерно 5 *в*, то обратный ток также начинает стремительно расти при увеличении напряжения, тем самым ставя практический предел обратного напряжения, которое выпрямитель может выдержать без потери выпрямительных свойств или нагрева до опасной температуры за счет выделения тепла при обратном токе. Критерием качества выпрямигеля, часто применяемым для характеристики его работы, является коэффициент выпрямления, определяемый как отношение прямого тока к обратному при некотором установленном напряжении. Величины коэффициентов выпрямления для различных контактных выпрямителей лежат в пределах от нескольких сотен до

1 млн. при 1 *в*. Для характеристики, представленной на рис. 5-2, коэффициент выпрямления при 1 *в* равен 2 000.

Некоторые особенности характеристик выпрямителей, подобных показанной на рис. 5-2, иногда можно представить более удобно, построив эту характеристику в координатах с логарифмическим масштабом тока и напряжения, как это сделано на рис. 5-3. Из рассмотрения физического механизма работы выпрямителя следует, как это будет ясно из последующих глав, что ток удобно принимать за независимую переменную. Поскольку двойные логарифмические координаты не имеют начальной точки, прямые и обратные характеристики выпрямителя построены отдельно в виде двух различных ветвей, сходящихся при очень малых токах (область *А* на графике). Помимо выявления деталей характеристик, теряющихся при построении в линейном масштабе, логарифмические координаты позволяют легко определить коэффициент выпрямления, прочитав его прямо на графике, и вдобавок имеют то преимущество, что наклон характеристики в каждой рабочей точке сразу дает величину показателя n, определяющего зависимость тока от напряжения в данной рабочей точке. Таким образом, если $I = \text{const} \cdot U^n$, то

$$\left.\begin{aligned} \lg I &= \lg \text{const} + n \lg U \\ \text{и} \qquad \frac{d(\lg U)}{d(\lg I)} &= \frac{1}{n}. \end{aligned}\right\} \qquad (5\text{-}1)$$

На характеристике прямого тока обычно имеется участок с практически постоянным показателем n, охватывающий один или два десятичных порядка изменения тока. Если ток в какой-либо области характеристики возрастает пропорционально приложенному напряжению (закон Ома), то логарифмическая характеристика в этой области имеет наклон 45°. Характеристика выпрямителя на рис. 5-3 приближается к линейной при очень малых токах обоих направлений (область *А*), при большом прямом токе (область *Г*) и при обратном токе средней величины (область *Д*).

Различные особенности характеристики на рис. 5-3 могут быть просто истолкованы с помощью эквивалентной схемы контактного выпрямителя, показанной на рис. 5-4. Выпрямительный контакт представлен парал-

лельно соединенными сопротивлениями барьерного слоя R и емкостью этого слоя C. Последняя представляет собой емкость конденсатора, образованного металлом, массой полупроводника и тонким запирающим слоем между ними, играющим роль диэлектрика. Сопротив-

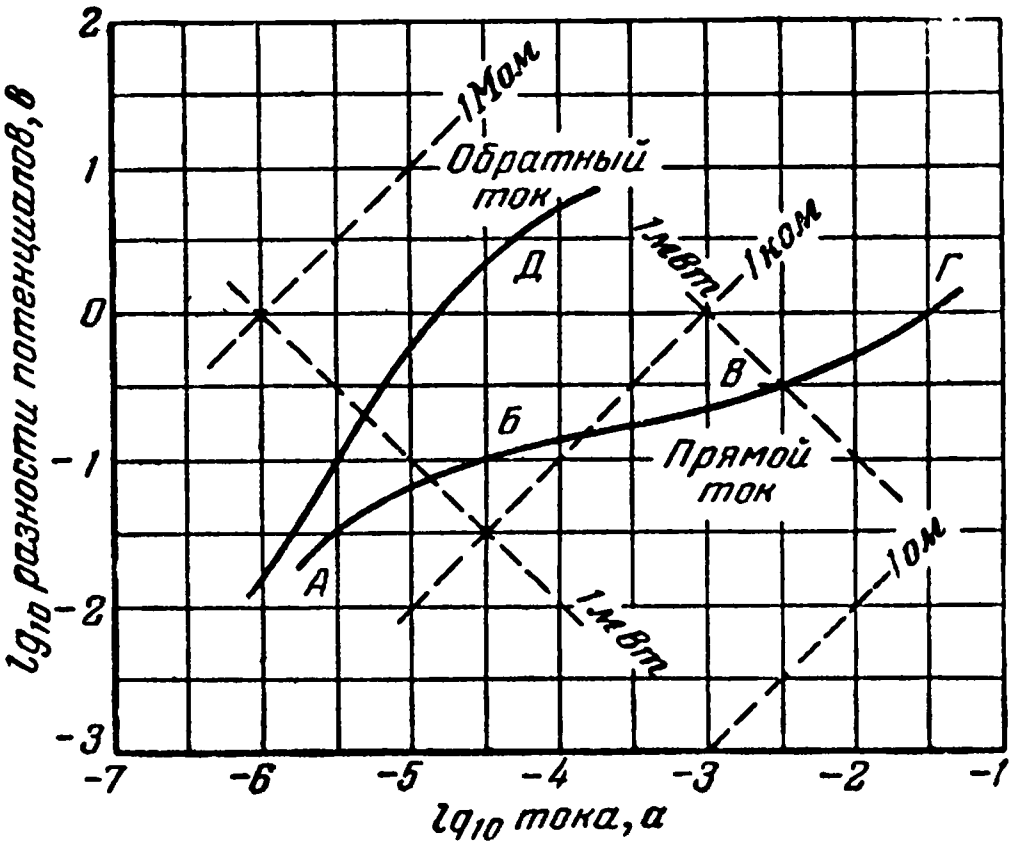


Рис. 5-3. Так выглядит характеристика меднозакисного выпрямителя, приведенная на рис 5-2, если построить ее в двойном логарифмическом масштабе.

ление R непостоянно, его величина зависит от направления и величины напряжения на слое. Величина C также зависит от напряжения. Последовательно включенное сопротивление r представляет собой омическое сопротивление тела полупроводника, соединенного последовательно с выпрямительным контактом. При низких напряжениях R во много раз больше, чем r, и характеристика в основ-

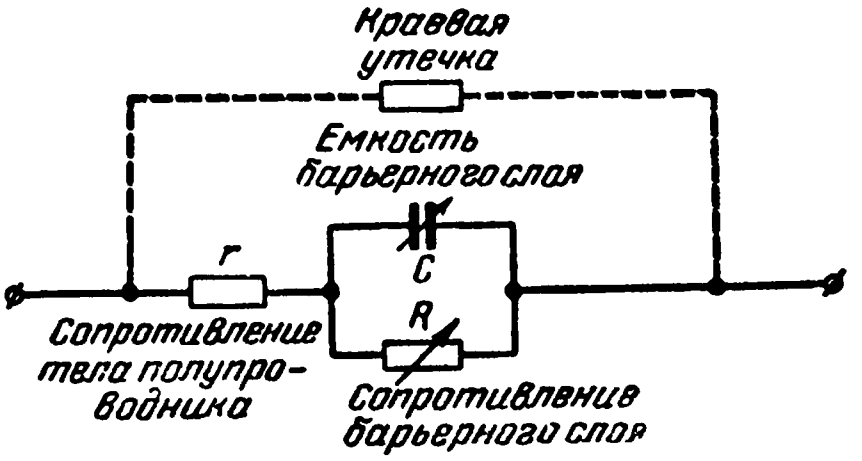


Рис 5-4 Эквивалентная электрическая схема выпрямительного элемента

ном определяется поведением R. Можно показать, что R имеет в основном характер омического сопротивления, если приложенное напряжение U в любом направлении меньше, чем kT/e, где k — постоянная Больцмана; T — абсолютная температура и e — заряд электрона.

При комнатной температуре величина kT/e равна примерно 0,025 *в*. Этот почти чисто омический характер R при малых напряжениях отображается существованием асимптоты характеристики в области A, наклоненной под углом 45°. Если повышать приложенное прямое напряжение, то R постепенно уменьшается, пока, наконец, при напряжении около 1 *в* R становится меньше r, имеющего омический характер и определяющего характеристику при более высоких напряжениях, вследствие чего в крайней правой части рис. 5-3 она стремится к асимптоте, наклоненной под углом 45° (область *Г*).

Омическое сопротивление r называется прямым ограничивающим сопротивлением выпрямителя. На практике стремятся свести r к минимуму, чтобы сохранить нелинейность характеристики в прямом направлении до возможно большего тока (характеристика не должна приближаться к идущей под углом 45° предельной линии, отображающей омическое сопротивление) с целью создать выпрямитель с большим коэффициентом выпрямления. Уменьшения r можно добиться, уменьшая толщину слоя полупроводника до тех пор, пока не возникнет опасность образования неоднородного слоя или нарушения его непрерывности; можно также до известных пределов уменьшить величину r понижением удельного сопротивления полупроводника. На рис. 5-5 показано семейство характеристик купроксного выпрямителя, соответствующих трем разным толщинам слоя закиси меди [Л. 1]. Можно видеть, что выпрямители с более тонким слоем закиси меди характеризуются бóльшим прямым током и более высоким коэффициентом выпрямления.

В некоторых контактных выпрямителях (особенно в купроксных) нужно принимать в расчет еще один элемент эквивалентной схемы, а именно проводимость краевой утечки, которая возникает из-за омической проводимости между контактными электродами через края полупроводникового слоя. Цепь этой утечки показана пунктиром на рис. 5-4. Краевая утечка нежелательна, поскольку она представляет собой обходный путь, парал-

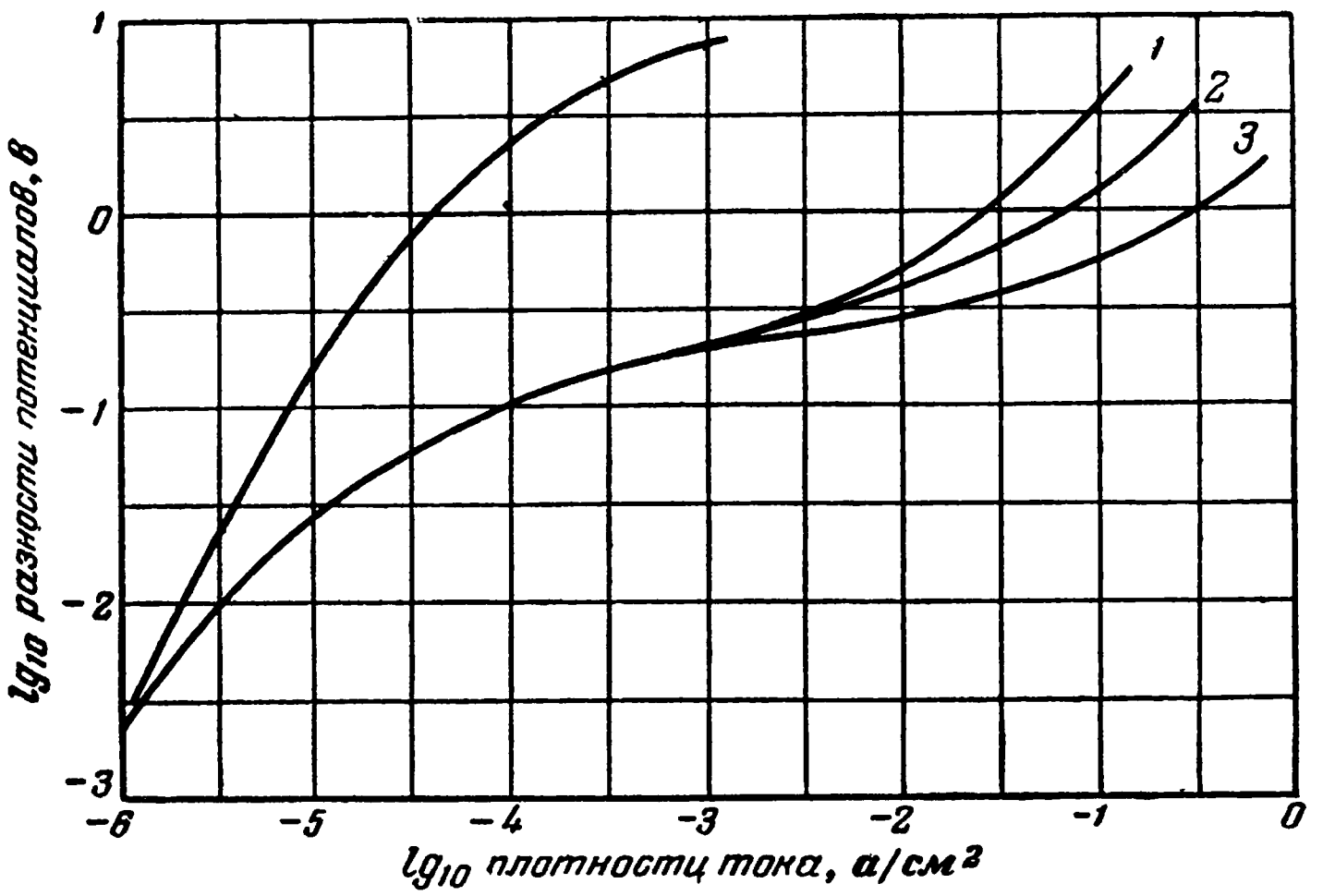


Рис. 5-5. Меднозакисный выпрямитель с более тонкой пленкой закиси меди обладает меньшим прямым предельным сопротивлением и пропускает больший прямой ток при напряжении 1 *в*. Толщина пленки меди: *1*—0,029 *см*; *2*—0,021 *см* и *3*—0,011 *см*

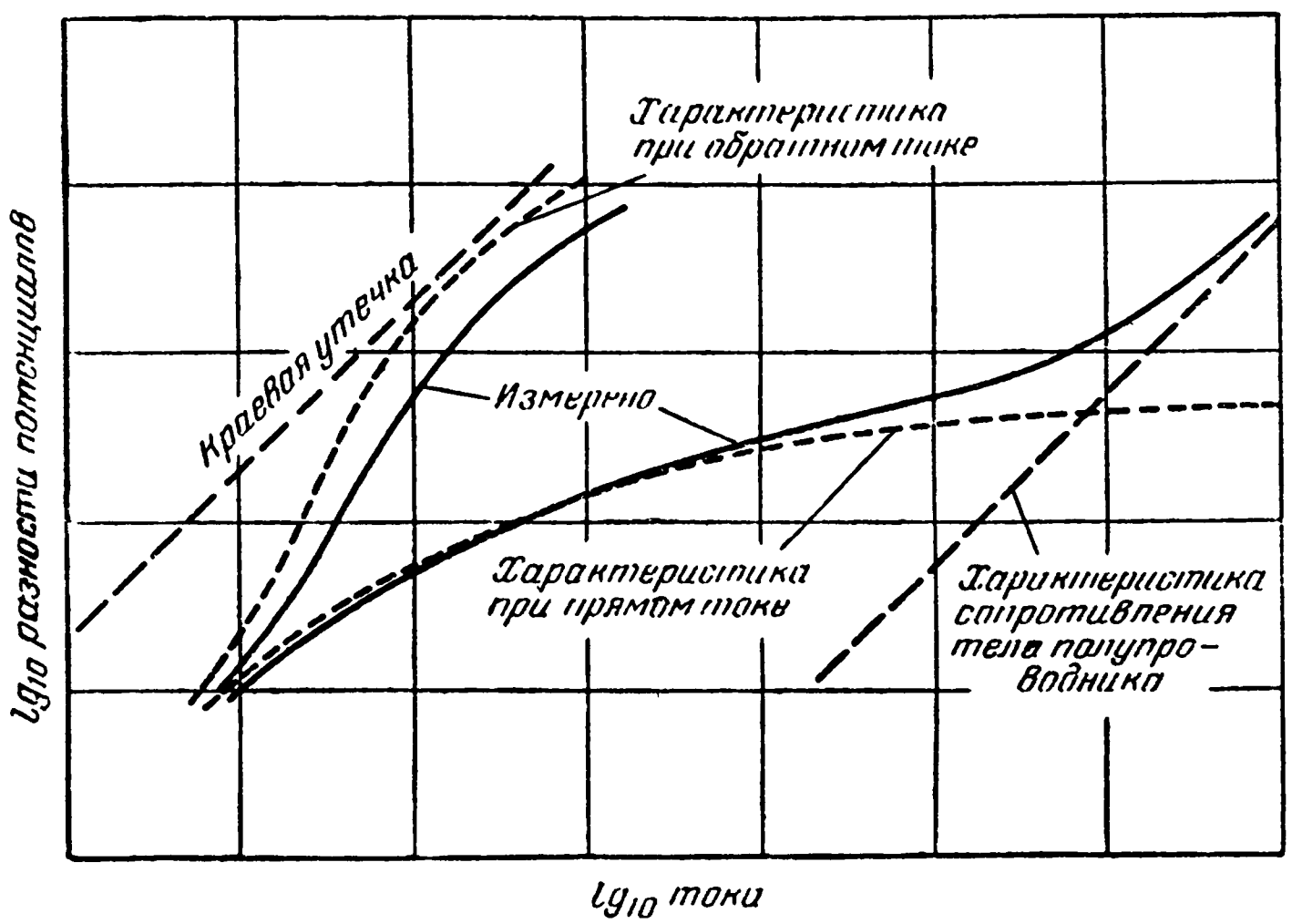


Рис 5-6. Синтез полной внешней измеренной характеристики выпрямительного элемента из характеристик отдельных элементов его эквивалентной схемы.

лельный нелинейной части выпрямителя; при изготовлении выпрямителей, в работе которых существенна утечка, принимаются меры для возможного уменьшения ее влияния. Краевая утечка является основной трудностью в выпрямительных элементах малого диаметра, у которых относительно велико отношение длины периметра к площади.

Легко показать, как реальная характеристика выпрямителя может быть составлена из частных характеристик элементов его эквивалентной схемы. На рис. 5-6 показан графический синтез сложной характеристики элемента, в котором значительную роль играет краевая утечка. Наоборот, измеренную характеристику выпрямителя можно разложить на частные характеристики барьерного слоя, предельного сопротивления в прямом направлении и сопротивления краевой утечки, если ее влияние значительно.

5-1-1. Купроксные и селеновые контактные выпрямители большой площади

Известно сравнительно большое число выпрямляющих контактных металл-полупроводниковых пар, но наибольшее внимание и наибольшие усилия разработчиков плоскостных выпрямителей были направлены на купроксные и селеновые приборы [Л. 2]. Из всех известных плоскостных приборов эти два вида, по-видимому, сочетают наилучшим образом следующие качества:

а) большую допустимую токовую нагрузку в прямом направлении на единицу контактной поверхности;

б) способность выдерживать сравнительно высокие обратные напряжения;

в) приемлемо высокие коэффициенты выпрямления;

г) достаточную стабильность и медленное старение.

Процесс производства купроксных выпрямителей [Л. 3 и 4] начинается вырубкой из медного листа шайб или заготовок в виде кружков наподобие монет. Размер шайб определяется нужной токовой нагрузкой или требуемым сопротивлением готового изделия. Заготовки затем химически очищаются (протравливаются) и помещаются шероховатой стороной вверх в печь при температуре примерно 1 000° С, где они окисляются с одной стороны, пока толщина слоя закиси меди не достигнет приблизительно 0,08 *мм*. Образовавшаяся красная закись

меди является полупроводником, а контакт между слоем закиси меди и металлом, из которого она образовалась, обладает выпрямительным свойством. Затем изделие отжигается при температуре около 500° С для повышения проводимости слоя закиси меди. Нагретые до температуры отжига заготовки закаливаются в холодной воде. Во время этой обработки окисленные заусеницы по краям шайбы скалываются, в результате чего уменьшается нежелательное влияние токов утечки через края готового элемента. Образующийся во время отжига тонкий слой

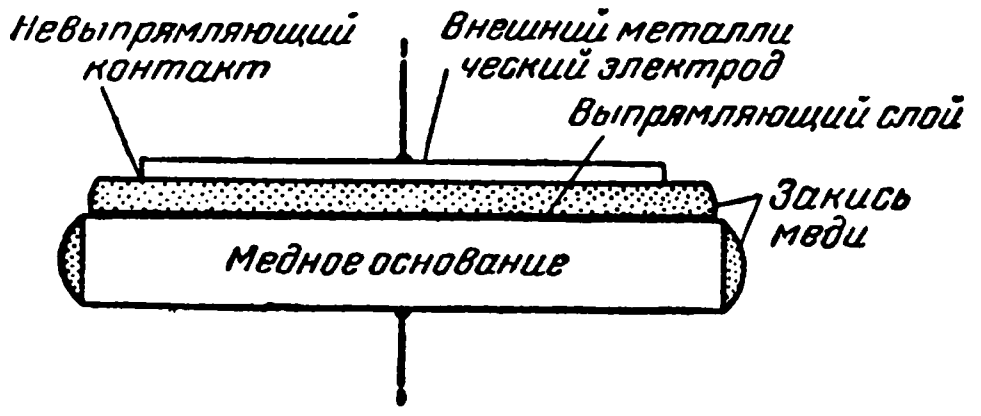


Рис. 5-7 Схематический разрез меднозакисного выпрямительного элемента. Прямой ток проходит при отрицательном напряжении на медном основании.

черной окиси меди стравливается в кислотной ванне, чем еще больше уменьшается проводимость краевой утечки. После промывки и просушки поверхность красной закиси меди слегка обрабатывается пескоструйным аппаратом и затем на нее наносится невыпрямляющий контактный слой, обычно путем вакуумного испарения серебра или золота. Готовый купроксный выпрямительный элемент показан в разрезе на рис. 5-7. На этом чертеже толщины слоев закиси меди и внешнего металлического электрода сильно преувеличены. На рис. 5-8 показана фотография двух собранных купроксных выпрямителей.

Рабочие части селенового выпрямителя [Л. 5 и 6] расположены на подложке из никеля, никелированного железа или алюминия, являющейся основой; диаметр подложки выбирается соответственно заданному рабочему току выпрямителя. На одну сторону подложки наносится слой селена путем наплавления при температуре немного выше его точки плавления (218° С) либо конденсацией паров селена в вакууме. Затем нанесенный слой подвергается термической обработке, переводящей селен из аморфной фазы в кристаллическую, в которой он об-

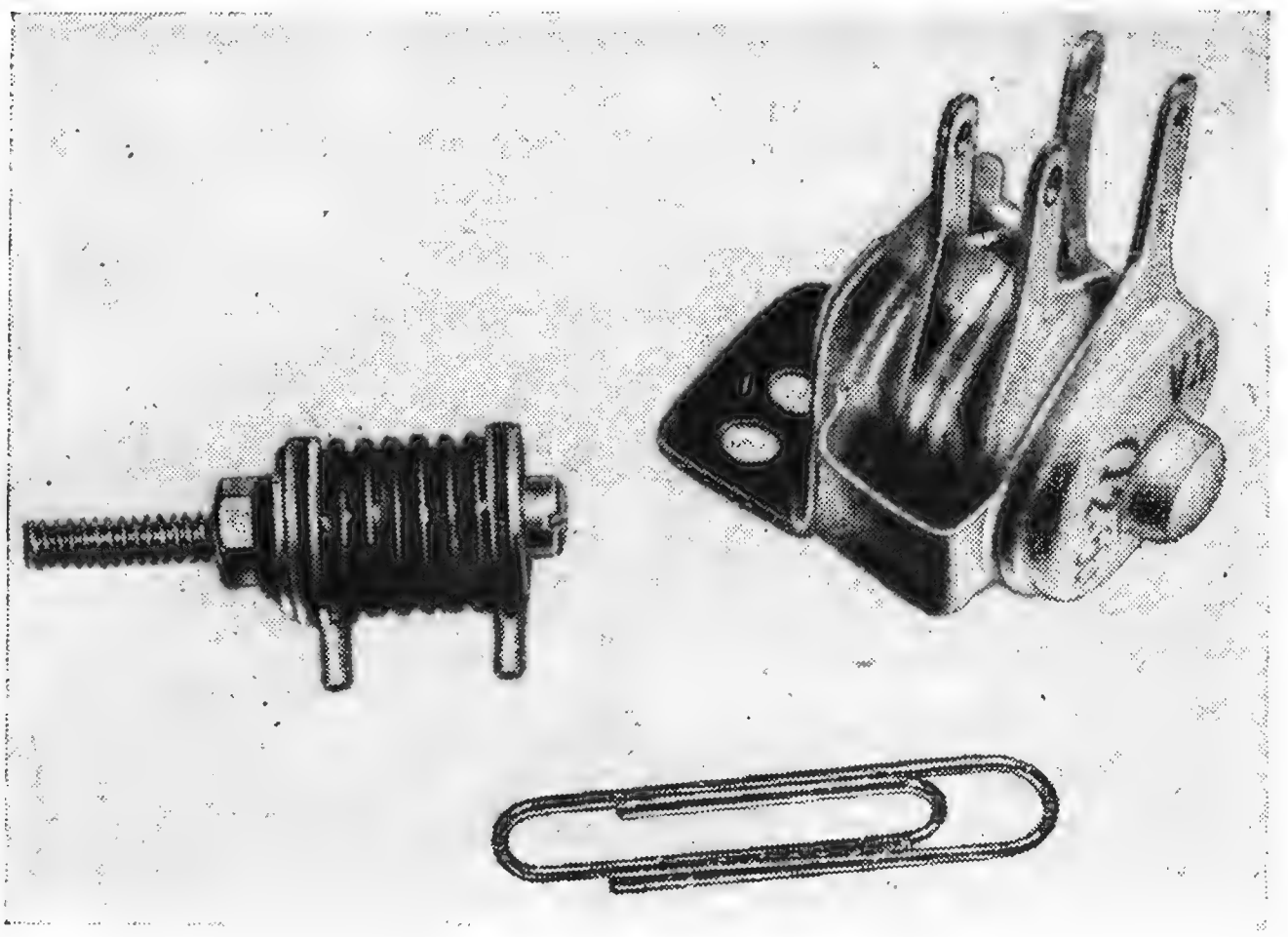


Рис 5-8. Слева — внешний вид однополупериодного меднозакисного выпрямителя, собранного из элементов диаметром 11 *мм*, обеспечивающего 50 *ма* выпрямленного тока при переменном напряжении 45 *в* на входе. Справа — мостовой выпрямитель, собранный из четырех элементов диаметром $\sim$5 *мм*, обеспечивающий выпрямленный ток 200 *ма* при напряжении 6 *в*.

наруживает свойства полупроводника. Наконец, на свободную поверхность селена наносится контактный слой напылением кадмия или сплава на кадмиевой основе.

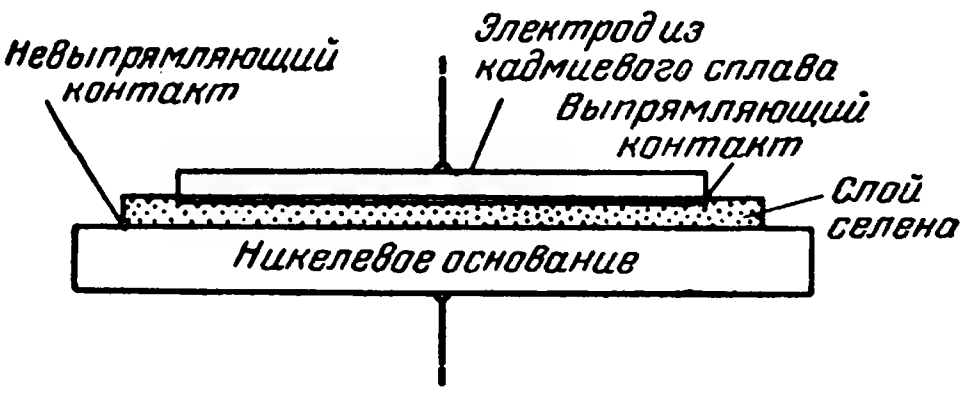


Рис. 5-9. Схематический разрез селенового выпрямительного элемента. Прямой ток проходит при положительном напряжении на никелевом электроде.

В селеновом выпрямителе выпрямляющим является контакт между селеном и этим сплавом, тогда как контакт между селеном и подложкой является в основном омическим и невыпрямляющим. На рис. 5-9 показан схема-

тический разрез селенового выпрямителя. Здесь также преувеличены толщины полупроводникового и внешнего контактного слоев. На рис. 5-10 приведена фотография двух собранных селеновых выпрямителей, выпускаемых промышленностью.

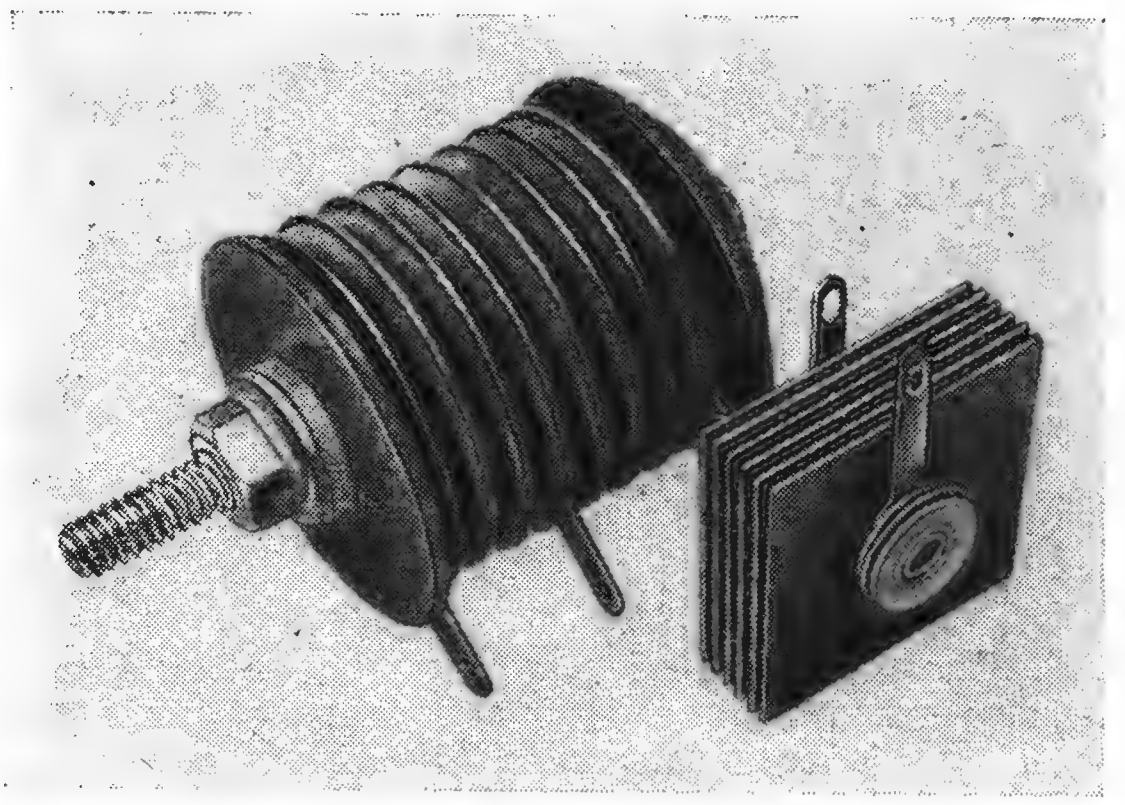


Рис. 5-10. Селеновый выпрямитель слева предназначен для работы в схеме двухполупериодного выпрямления с отводом от средней точки. Выпрямитель справа предназначен для работы в однополупериодной схеме. Оба выпрямителя имеют эффективную площадь выпрямляющего контакта ~ 8 *см*^{2} и могут обеспечить до 300 *ма* выпрямленного тока.

Окончательной стадией изготовления селенового выпрямителя является электрическая обработка, называемая формованием. При этом процессе к выпрямителю прикладывается обратное напряжение и пропускается обратный ток. В результате этого возрастает обратное сопротивление выпрямителя. Формование производится при постепенном повышении напряжения, иногда в течение нескольких часов, после чего обратное напряжение может повыситься в 100 и более раз, достигнув величины от 10 до 20 *в*. Характеристики прямого и обратного тока в области малых напряжений при формовании существенно не изменяются.

До сих пор еще не найдено способов стабилизации характеристик селенового выпрямителя в отформован-

ном состоянии. Если непрерывная работа выпрямителя предохраняет его от потери приобретенных при формовании свойств, то пребывание в бездействии в течение нескольких дней или недель обычно приводит к восстановлению обратных характеристик в том виде, какой они имели до формовки, вследствие чего перед очередным включением необходимо снова формовать выпрямитель для безопасного включения при полном обратном

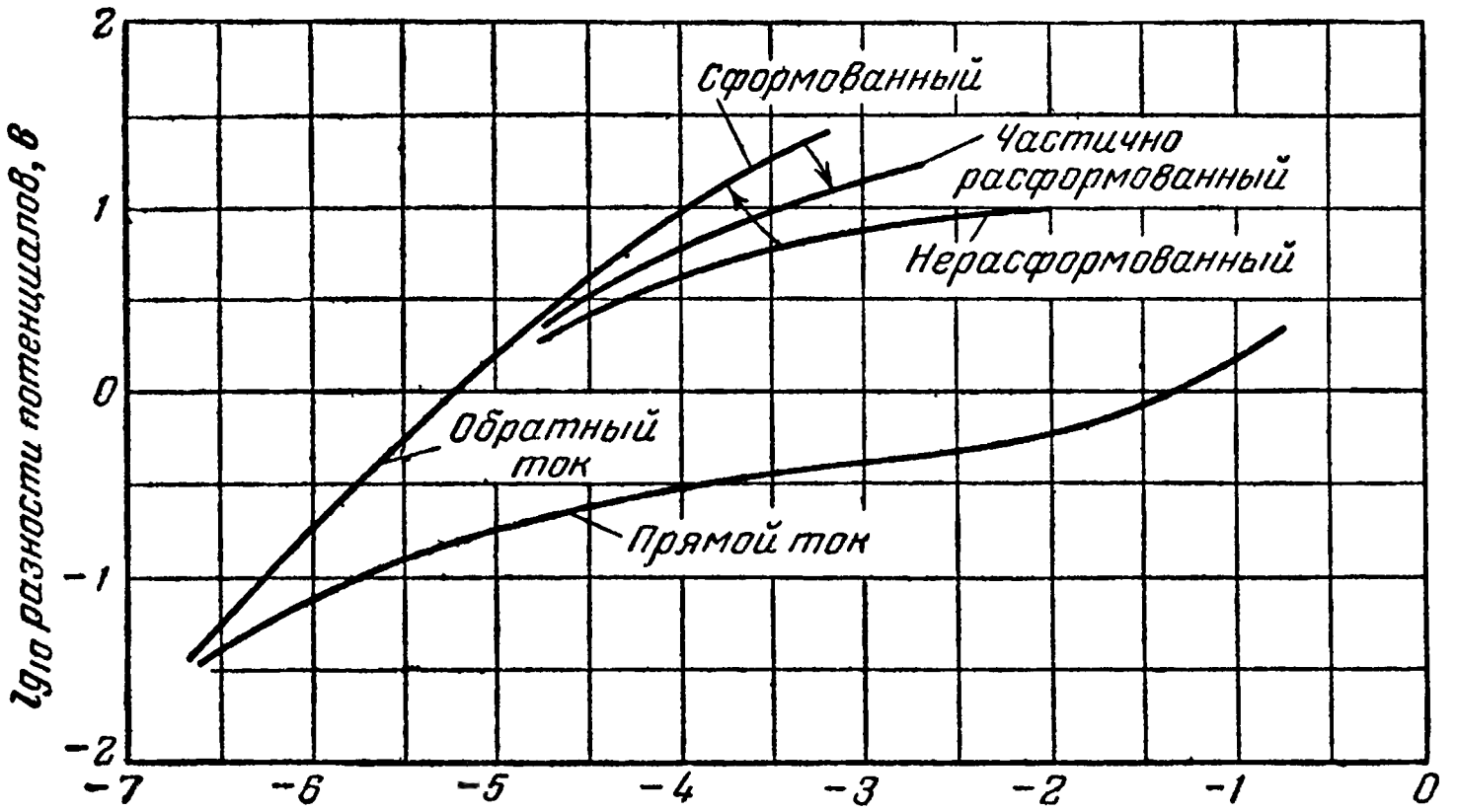


Рис. 5-11. Зависимость характеристик обратного тока селенового выпрямителя от степени формовки и предшествующего срока службы.

напряжении. Из-за свойства самопроизвольной расформовки при бездействии селеновые выпрямители мало пригодны при редкой или непостоянной работе. На рис. 5-11 показаны сравнительные характеристики селенового выпрямителя в неотформованном, отформованном и частично расформованном состояниях.

Как закись меди, так и селен являются полупроводниками *p*-типа; следовательно, в этих двух типах выпрямителей полупроводник положителен по отношению к контактному слою при прямом токе. На рис. 5-12 приведены для сравнения характеристики купроксного и селенового выпрямителей с единичной контактной площадью. Нужно иметь в виду, что изготовители как купроксных, так и селеновых выпрямителей все время меняют детали технологии производственного процесса,

вследствие чего «стандартные» характеристики часто изменяются.

Купроксные и селеновые элементы действуют в качестве выпрямителей так, что во время полупериода прямого тока среднее падение напряжения на выпрямителе составляет примерно 1 *в*. Из рис. 5-12 можно видеть, что при напряжении 1 *в* купроксный выпрямитель пропу-

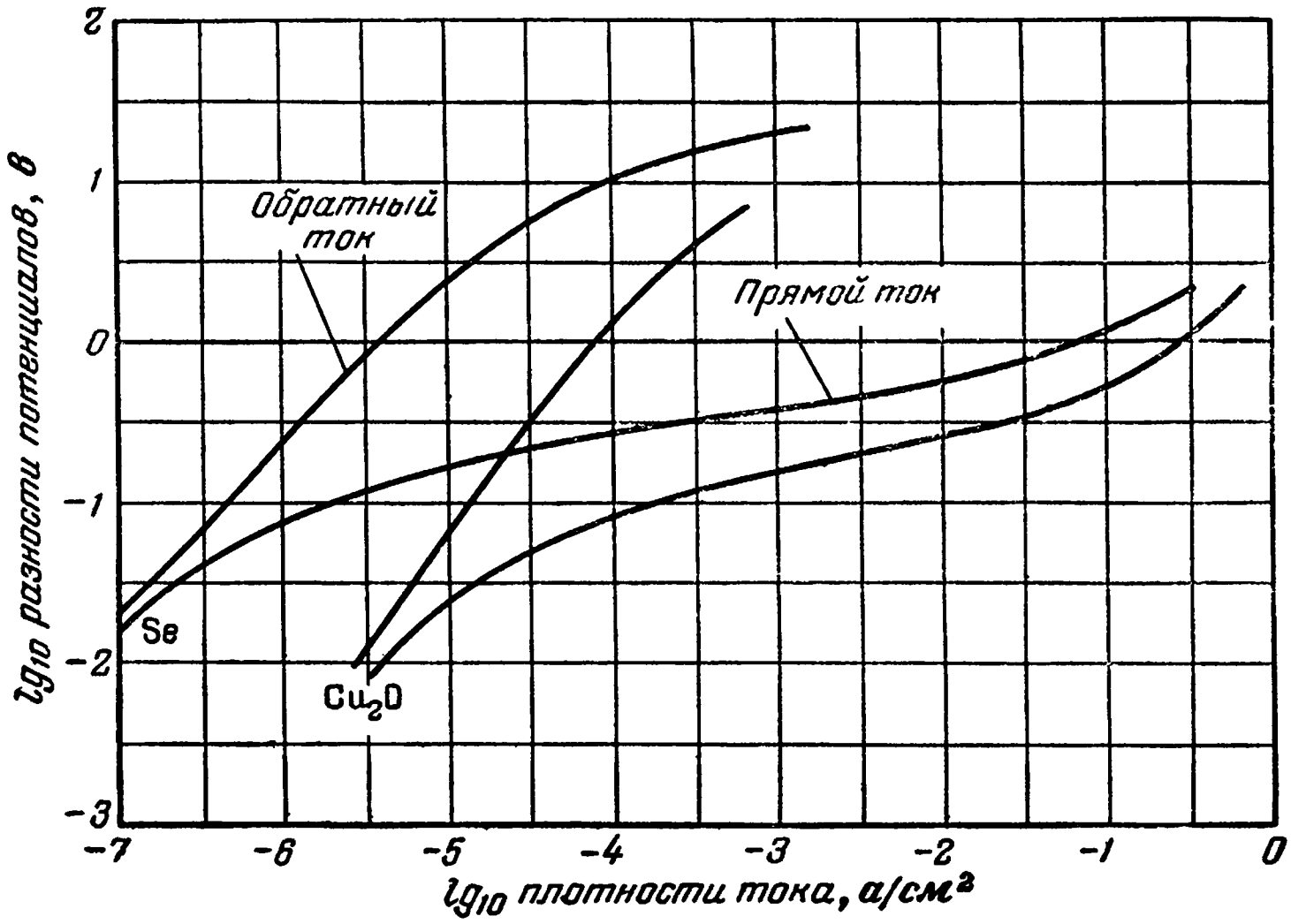


Рис. 5-12. Характеристики селенового и меднозакисного выпрямителей, пересчитанные для сравнения к единице площади контактной поверхности.

скает несколько больший ток на единицу контактной поверхности, чем селеновый. Следовательно, при заданном выпрямленном токе размеры купроксного выпрямителя меньше, чем селенового. С другой стороны, селеновый выпрямитель может выдержать максимальное обратное напряжение от 20 до 30 *в* на элемент, тогда как купроксный выпрямительный элемент не выдерживает напряжения более 6—8 *в*. Если желательно получить выпрямленный ток при сравнительно высоком напряжении, селеновые выпрямители могут дать бóльшую экономию места и веса, поскольку потребуется включать последовательно меньшее число элементов, чтобы выдержать заданное максимальное обратное напряжение.

Купроксные выпрямители выпускаются промышленностью нескольких размеров, начиная от шайб диаметром 1,6 *мм* и кончая пластинами с площадью порядка 300 $см^2$. Прямой ток, пропускаемый этими выпрямителями при напряжении 1 *в*, находится в пределах от 2 *ма* для наименьших до 40—50 *а* для наибольших элементов.

Селеновые выпрямители изготовляются разных размеров от шайб диаметром 0,8 *мм* до пластин с площадью примерно 200 $см^2$, рассчитанных на выпрямленный ток от 5 до 10 *а* при максимальном обратном напряжении 20 *в* на пластину.

5-1-2. Плоскостные выпрямители с *p-n* переходом

Как уже было отмечено, выпрямление может происходить не только в контактах между металлами и полупроводниками, но также и в контактах между различными полупроводниками. Применяемые при этом полупроводники могут быть совершенно различными хими-

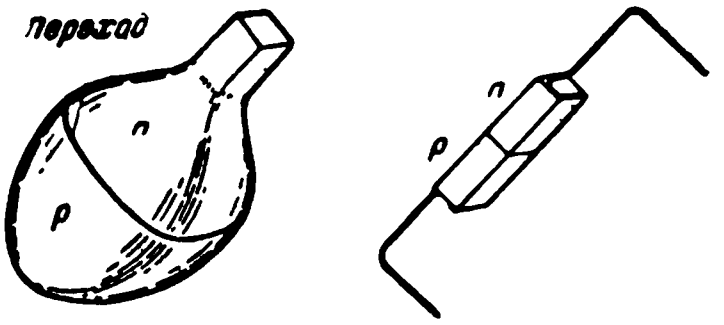


Рис. 5-13. Слева — выращенный кристалл германия с *p-n* переходом, справа — выпрямляющий стержень с *p-n* переходом, вырезанный из кристалла.

чески, но могут быть также просто разными участками одного и того же полупроводника, отличающимися только типами проводимости.

Наиболее известным примером таких выпрямителей является германиевый выпрямитель с *p-n* переходом [Л. 7]. Для его изготовления приготовляется слиток германия, одна часть которого имеет проводимость *n*-типа, а другая — *p*-типа. Как это достигается с помощью надлежащего выбора и распределения примесей в полупроводнике, описано в гл. 25. Весь слиток представляет собой монокристалл, на половине длины которого перпендикулярно оси его роста расположена граница между областями с проводимостями *p*-типа и *n*-типа (рис. 5-13).

Этот слиток разрезается на множество прямоугольных стержней, ориентированных относительно поверхности перехода так, что она делит каждый стержень на две части: один конец—*n*-типа, другой—*p*-типа. Эта граница, называемая *p-n* переходом, является выпрямляющим участком всего элемента; контакты подводящих проводов делаются невыпрямляющими. На рис. 5-14 приведена фо-

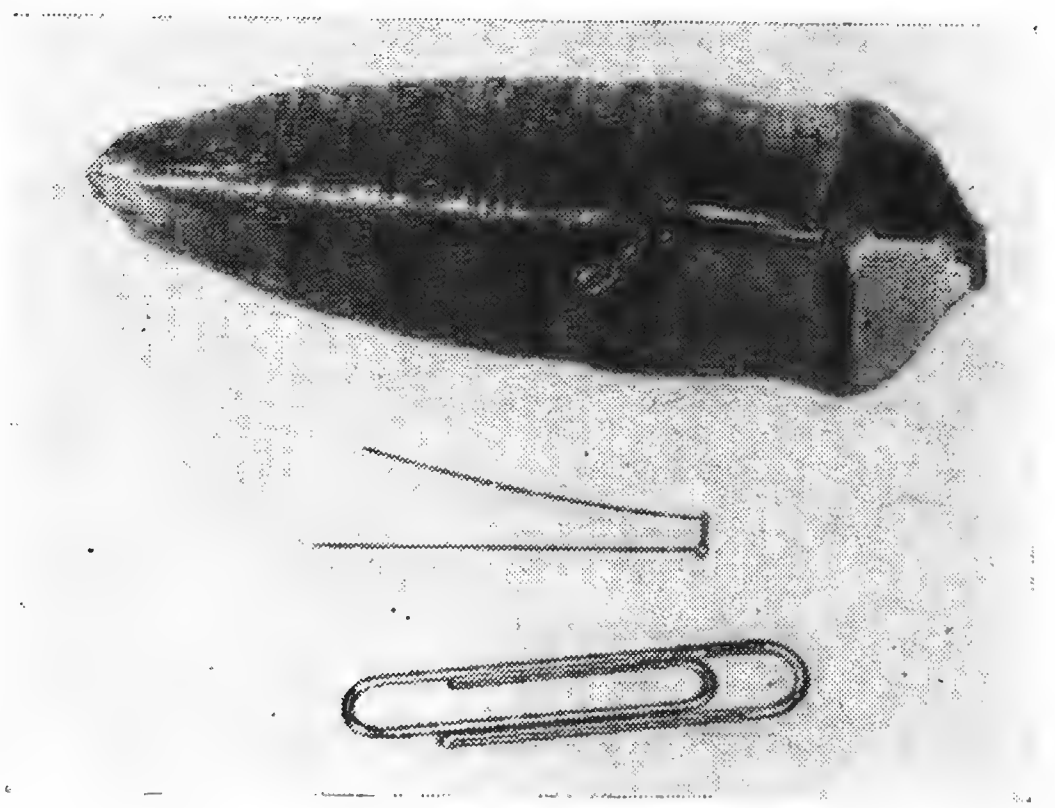


Рис. 5-14. Внизу выпрямляющий стерженек с *p-n* переходом, имеющий проволочные выводы длиной около 3 *мм* и площадью сечения около 1 *мм*^{2}. Стержень вырезан из выращенного кристалла, подобного показанному наверху.

тография выращенного германиевого слитка и вырезанного из него стержня для выпрямителя с *p-n* переходом.

Площадь поперечного сечения стержня, определяющая площадь выпрямляющего перехода, выбирается соответственно требуемому рабочему току выпрямителя. Наиболее часто применяемый размер стержня — около 1 *мм*^{2} в поперечном сечении и около 3 *мм* длины. Выпускаются выпрямительные элементы бóльших размеров — до 1 *см*^{2} в сечении, но они не могут работать при высоких плотностях прямого тока из-за трудности отвода выделяющегося тепла при относительно больших размерах элемента. На каждом стержне создаются электрические невыпрямляющие контакты, к которым присоединяют проволочные выводы. Затем элемент заключается в защитную оболоч-

ку для защиты от механических повреждений при эксплуатации, изоляции поверхности стержня от влаги и защиты переходной зоны от света.

Выпрямительная характеристика типичного германиевого выпрямителя с *p-n* переходом показана на рис. 5-15. Сравнение этой характеристики с характеристиками меднозакисных и селеновых выпрямителей показывает, что

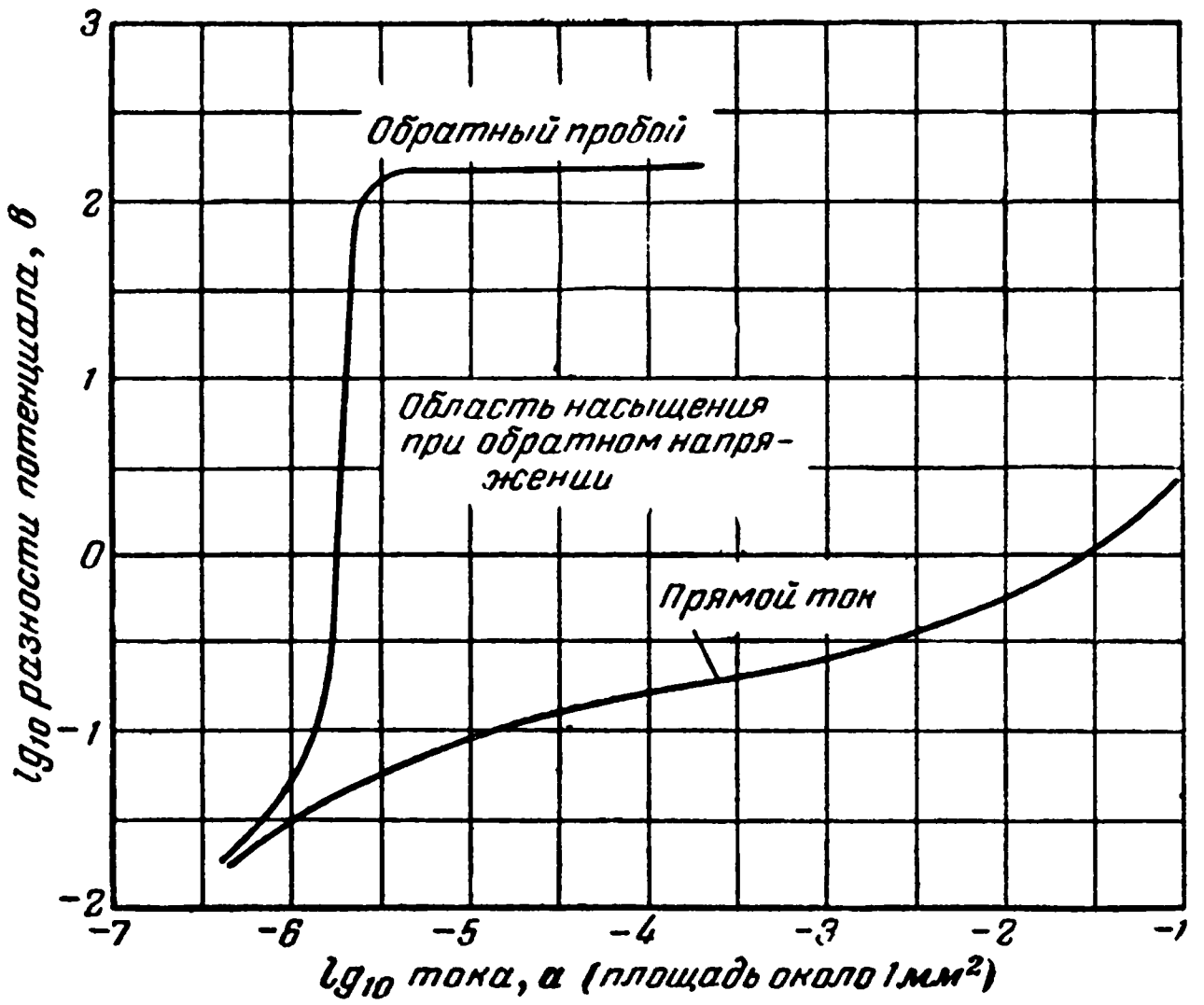


Рис 5-15. Обратите внимание на участки насыщения и пробоя на характеристике обратного тока выпрямителя с *p-n* переходом.

прибор с *p-n* переходом пропускает примерно в 10 раз больший предельный прямой ток на единицу площади, чем любой другой тип выпрямителя, тогда как обратный ток на единицу площади при напряжении 1 *в* сравним с таковым для меднозакисного выпрямителя. Заметим, однако, что описанный выпрямитель с переходом выдерживает без пробоя обратное напряжение примерно до 150 *в*. Таким образом, с помощью одного элемента можно создать высоковольтный выпрямитель или триггер для сигналов большой амплитуды.

Часть характеристики выпрямителя с *p-n* переходом, соответствующая прямому току, имеет те же особенности, которые характерны для выпрямителей других типов;

обратный же ток остается примерно постоянным в области напряжений от 0,1 *в* до ~150 *в*. Это насыщение обратного тока не является чем-то аномальным или необычным, а оказывается именно той характерной чертой, которую можно ожидать теоретически от «чистого» выпрямителя. Только после создания выпрямителя с *p-n* переходом впервые оказалось возможным получить такую идеальную характеристику обратного тока [Л. 8], хотя она была предсказана теоретически за несколько лет до этого [Л. 9, 10 и 11]. Насыщение обратного тока не наблюдается в сколько-нибудь заметной степени на характеристиках меднозакисных и селеновых выпрямителей или точечных диодов по причинам, которые разъяснены ниже. Ток в указанной области напряжений называется обратным током насыщения элемента $I_{нас}$. Его величина у типичных германиевых выпрямителей с *p-n* переходом обычно составляет около 1—2 *мка* на 1 *мм*^{2} площади перехода.

Теоретический вывод уравнения характеристики выпрямителя с переходом приводит к следующему выражению:

$$I = I_{нас}\left(1 - e^{eU_a/kT}\right) \; a/см^2, \qquad (5\text{-}2)$$

где $I_{нас}$ определено выше; e — заряд электрона ($1{,}6 \times 10^{-19}$ k); k — постоянная Больцмана ($1{,}38 \cdot 10^{-23}$ джоуль на градус); T—абсолютная температура и U_a—приложенное напряжение, принятое за положительное для прямого направления.

При первом взгляде на уравнение (5-2) может показаться, что под действием прямого напряжения создается отрицательный ток. Однако, как следует из определения, $I_{нас}$ является существенно отрицательным, так что никаких затруднений со знаками и направлением тока не возникает. Это уравнение дает форму как обратной, так и прямой ветвей характеристики вплоть до наступления пробоя в обратном направлении и до приближения к предельному прямому сопротивлению в прямом направлении. Особенностью, предопределяющей интересные возможности в физике полупроводниковых приборов, является то, что $I_{нас}$ может быть рассчитан с помощью основных физических величин и свойств полупроводника по обе стороны от поверхности перехода. Этот подход рассмотрен более подробно в гл. 19.

Заметим также, что при повышении обратного напряжения характеристика выпрямителя резко изгибается, переходя из области насыщения в область, где ток при дальнейшем повышении напряжения очень быстро возрастает. Эта область называется областью обратного пробоя, а напряжение, соответствующее перегибу характеристики, — пробивным напряжением. Пробивное напряжение выпрямителя с *p-n* переходом находится под контролем разработчика; путем выбора исходного материала, имеющего надлежащие проводимости, можно обеспечить любое пробивное напряжение в пределах от 1 до более чем 1 000 *в*. Таким образом, характеристика на рис. 5-15 является только частным случаем из множества сильно отличающихся индивидуальных характеристик для различных вырямителей с *p-n* переходами.

Наличие на характеристиках выпрямителя с *p-n* переходом излома, соответствующего обратному пробою, используется в схемах ограничителей формы сигнала, защиты от перегрузок, ограничителей напряжения и импульсных амплитудных селекторов. Физическая сторона явления пробоя очень хорошо выяснена и изложена более подробно в гл. 19.

5-1-3. Силовые плоскостные диоды с переменной проводимостью

В предыдущих параграфах мы рассмотрели «обычные» плоскостные диоды, использовав в качестве примеров меднозакисный и селеновый контактные выпрямители и германиевый выпрямитель с *p-n* переходом. Теперь мы рассмотрим плоскостной прибор нового вида, который способен пропускать в сотни раз бóльшие прямые токи при сравнимых напряжениях и действующих площадях. Этот новый прибор называется выпрямителем с переменной проводимостью по причинам, которые будут ясны из дальнейшего изложения.

Лучшим и в настоящее время почти единственным примером диода с переменной проводимостью является кремниевый силовой диод [Л. 12], устроенный следующим образом. Пластина кремния *p*-типа толщиной около 0,13 *мм* подвергается с одной стороны воздействию паров фосфора при температуре в несколько сотен градусов. В этих условиях в поверхностный слой кремниевой пластины диффундирует достаточное количество фосфора, чтобы перевести этот слой на глубину около 0,01 *мм*

из p-типа в n-тип с высокой удельной проводимостью. Поверхность раздела между исходным материалом p-типа и преобразованным n-типом представляет собой p-n переход, обусловливающий выпрямительные свойства всего элемента. Другая сторона кремниевой пластины подвергается термообработке в атмосфере, содержащей пары бора, который подобным же образом диффундирует в поверхность на небольшую глубину, преобразуя подповерхностный слой кремния из обычного p-типа в p-тип

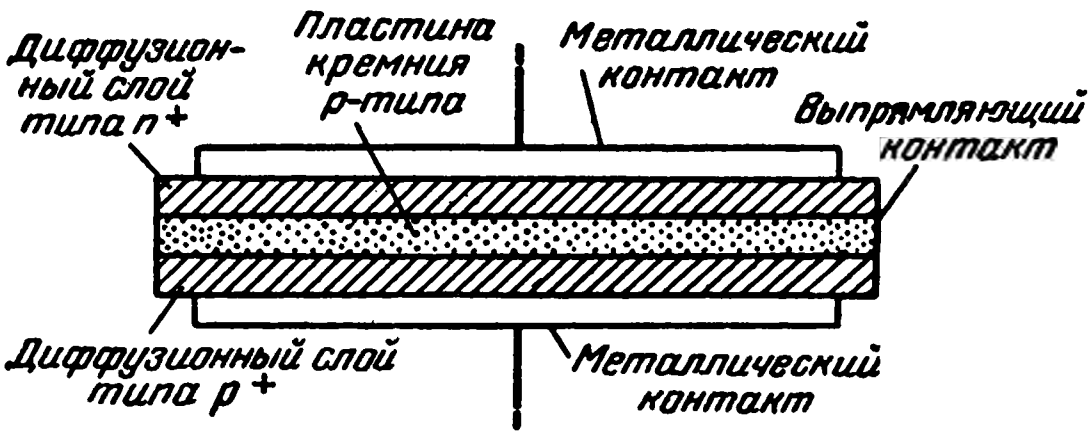


Рис 5-16. Схематический разрез кремниевого выпрямителя с переменной проводимостью.

с весьма высокой удельной электропроводностью (p^+) *. Поверхностные слои на обеих сторонах пластины имеют почти такую же электропроводность, как металлы, и на этих двух поверхностях сравнительно легко сделать контакты (что является очень трудной задачей при «обычном» кремнии). Промежуточные переходы p-p^+ не влияют на электрические свойства диода.

На рис. 5-16 показан схематический разрез такого диода, а на рис. 5-17 — фотография опытной модели диода на ток 10 *а* в защитной оболочке. Выпрямительный элемент, находящийся внутри, имеет рабочую поверхность 2,5×2,5 *мм* и одной стороной припаян к части кожуха, имеющей винтовую резьбу. При монтаже эта часть привинчивается к шасси для лучшего охлаждения, так как в диоде при работе на полную нагрузку 10 *а* выделяется от 5 до 10 *вт*.

Характеристики диода с переменной проводимостью

---

* Символы p^-, p и p^+ используются для обозначения материала p-типа с очень малой концентрацией акцепторной примеси, обычной концентрацией акцепторной примеси и очень большой концентрацией акцепторной примеси соответственно.

показаны на рис. 5-18. Токи пересчитаны на 1 *см*² рабочей площади для сравнения с данными для меднозакисного и селенового диодов, приведенными на рис. 5-12. Видно, что прибор с переменной проводимостью имеет

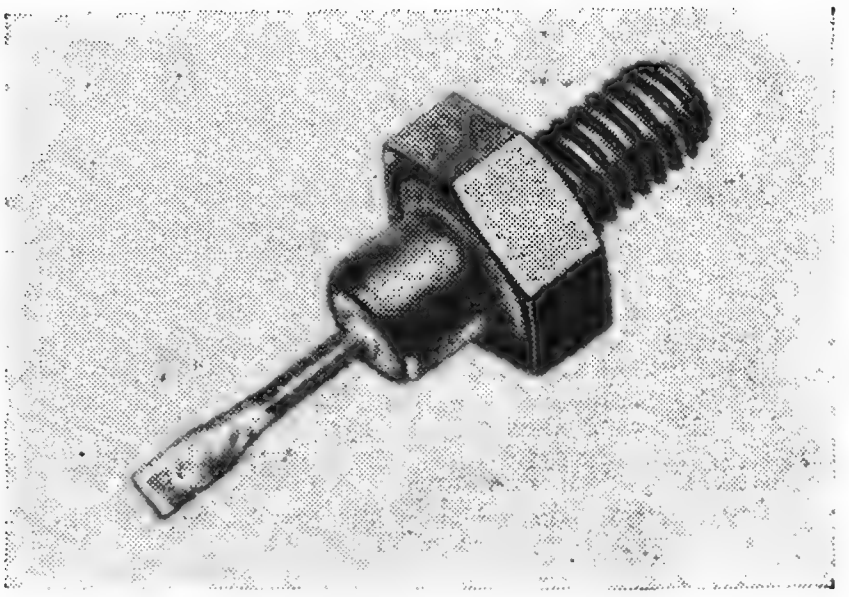


Рис 5-17 Опытный образец кремниевого выпрямителя с переменной проводимостью общей длиной около 32 *мм*, рассчитанный на 10 *а* среднего выпрямленного тока на элемент.

гораздо бóльшую прямую проводимость (например, при напряжении 1 *в*), чем «обычные» типы выпрямителей, тогда как его ток обратной утечки (при напряжении 1 *в*) не больше, чем у других типов. Можно видеть также, что

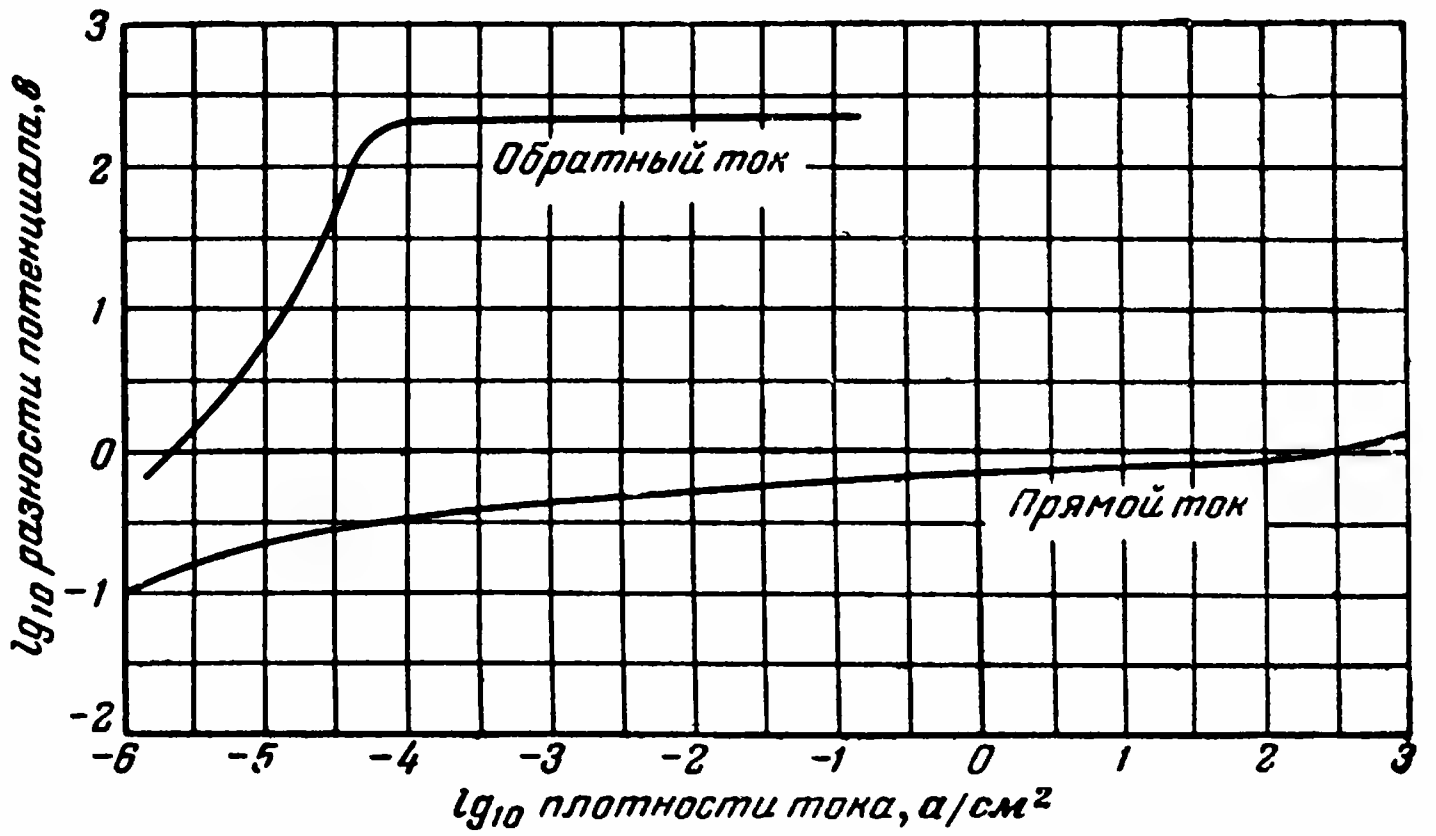


Рис. 5-18 Характеристика кремниевого выпрямителя с переменной проводимостью и необычайно высокой плотностью прямого тока при напряжении 1 *в*.

кремниевый диод выдерживает обратное напряжение в сотни вольт. Таким образом, по допустимому обратному напряжению один кремниевый диод может заменить несколько включенных последовательно меднозакисных или селеновых элементов. На основании имеющихся данных можно полагать, что эти приборы произведут революцию в области силовых выпрямителей.

Детальное рассмотрение физики работы диода с переменной проводимостью дано в гл. 19, а сейчас мы забежим вперед и дадим некоторое понятие о причинах необычайной проводимости этого прибора в прямом направлении. В обычных диодах уделяется мало внимания тому, является ли прямой ток потоком основных носителей из полупроводника в металлический контакт или же потоком неосновных носителей из контакта в полупроводник. При должном понимании физики работы прибора и надлежащем выборе технологии можно сделать выпрямительный контакт так, чтобы прямой ток создавался главным образом *неосновными* носителями, входящими в полупроводник. В кремниевом диоде с переменной проводимостью прямой ток является в основном потоком электронов из поверхностного слоя n-типа во внутреннюю область p-типа. Таким образом, прямой ток в этом приборе вводит в слой p-типа *избыточные* подвижные заряды. Чтобы при этом не возникало большого пространственного заряда в полупроводнике за счет этих неосновных носителей, должен иметь место соответствующий поток *избыточных* основных носителей с другого электрода, в нашем случае из p^+ слоя. Это внесение избыточных зарядов в слой полупроводника повышает его проводимость. В кремниевом диоде слой полупроводника p-типа при прохождении большого прямого тока приобретает проводимость, почти равную проводимости металла. Таким образом, предельное последовательное сопротивление может понизиться до сотых или тысячных долей его величины при малом токе, и диод при этих напряжениях будет пропускать соответственно больший прямой ток. Поскольку удельная проводимость полупроводника оказывается зависящей от прямого тока, такой прибор называется диодом с переменной проводимостью.

Мы не предполагаем, что это объяснение удовлетворит настоящего исследователя. Ведь ничего не говорится о роли, которую играет сильно насыщенный примесью

слой *n*-типа, почему этот слой играет роль контакта, инъектирующего неосновные носители, или почему обычные выпрямители не обнаруживают эффекта переменной проводимости в степени, достаточной для практического использования. Но, возможно, этот «взгляд за занавес» удовлетворит любопытство читателя на некоторое время, пока мы не рассмотрим этот вопрос более полно.

5-2. СИММЕТРИЧНЫЕ ВАРИСТОРЫ — ТИРИТ И КАРБИД КРЕМНИЯ

Симметричные варисторы имеют нелинейные характеристики симметричные при приложении напряжений различных знаков. Полупроводниковый материал, обнаруживающий такую нелинейность в степени, достаточной

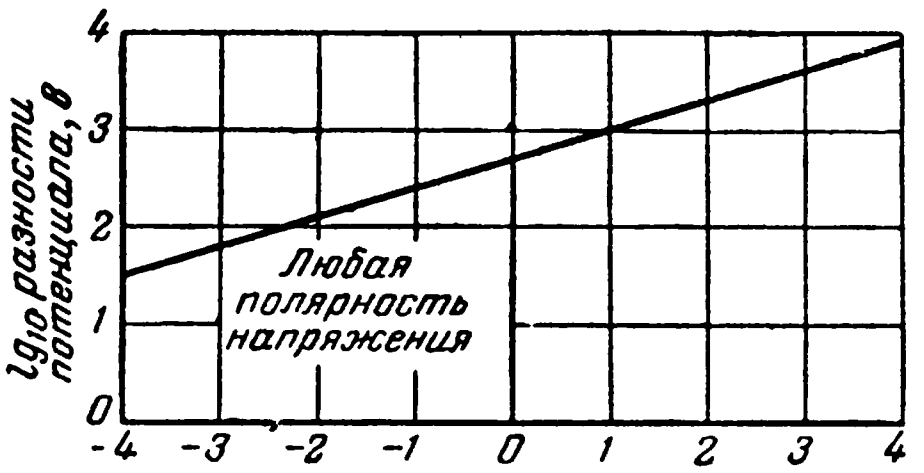


Рис. 5-19. Логарифмическая характеристика тиритового диска диаметром ~ 15 *см* и толщиной ~ 19 *мм*.

для практического использования, был разработан в 1930 г. и получил фирменное название тирит. Этот материал первоначально использовался для изготовления грозозащитных предохранителей [Л. 13] и с тех пор нашел много применений в области защиты от электрических перегрузок. На рис. 5-19 приведены логарифмические характеристики тиритового грозового предохранителя, имеющего форму диска диаметром ~ 15 *см* и толщиной 19 *мм* с металлическими электродами на противоположных сторонах. Характер нелинейности тирита подобен характеру нелинейности характеристики контактного выпрямителя при прямом токе, так как повышение напряжения на приборе вызывает много больший относительный прирост тока через него. Тиритовые диски изготовляются разных размеров и толщин, чтобы удовлетво-

рить самым различным требованиям, поскольку необходимы нелинейные схемные элементы на различные напряжения и токи. Никаких деталей технологии изготовления тирита не опубликовано.

Производство нелинейных элементов из карбида кремния начинается с обработки карбида кремния (SiC), который в большом количестве применяется в производстве абразивов. Для изготовления симметричных нелинейных элементов [Л. 14, 15 и 16] карбид кремния дробится в крупный порошок, смешивается с керамическим связующим веществом и небольшим количеством графита, смачивается и прессуется в диски, форма и размер которых определяются желательными характеристиками готовых элементов. Затем спрессованные заготовки спекаются в обжигательной печи, откуда они выходят в виде моно-

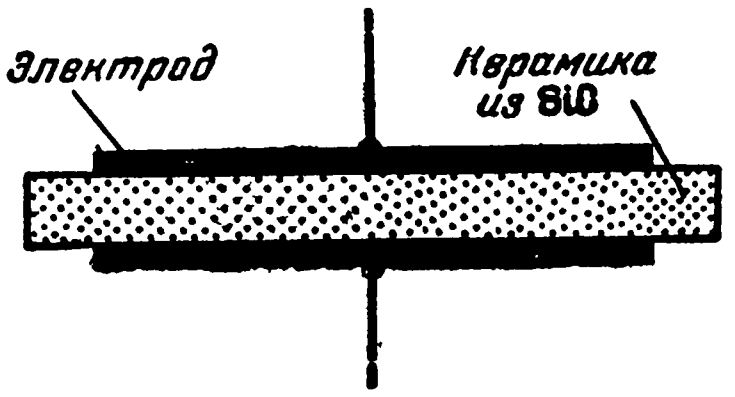


Рис. 5-20. Схематический разрез нелинейного элемента из карбида кремния, состоящего из сплавленной смеси карбида кремния, графита и керамики между металлическими электродами.

литных твердых дисков, в которых зерна карбида кремния прочно связаны керамической массой. Затем на диски наносятся электроды напылением металла или сплава или путем покрытия металлическими пастами с последующим восстановлением на них металла. На рис. 5-20 показан схематический разрез диска карбида кремния.

Электрические характеристики нелинейных элементов из карбида кремния зависят не только от их размеров и формы, но и от количества добавленного графита и керамики, размера и формы зерен карбида кремния и состава атмосферы в обжигательной печи.

Элементы из карбида кремния изготовляются в виде дисков и шайб диаметром 19 и 36 *мм* и толщиной от 0,75 до 3,2 *мм*. На рис. 5-21 приведены характеристики элемента из карбида кремния диаметром 19 *мм*, показывающие пределы, в которых можно менять параметры элемента только за счет режима термообработки. Эти характеристики симметричны и не зависят от полярности приложенного напряжения.

Причиной нелинейности элементов из карбида кремния является выпрямительный эффект, возникающий в точках соприкосновения зерен карбида кремния внутри элемента. Предполагается, что в зернах карбида при термообработке возникают области *p*- и *n*-типа, или области, проводимость которых (*p*- или *n*-типа) различается количественно; соприкосновение таких различных об-

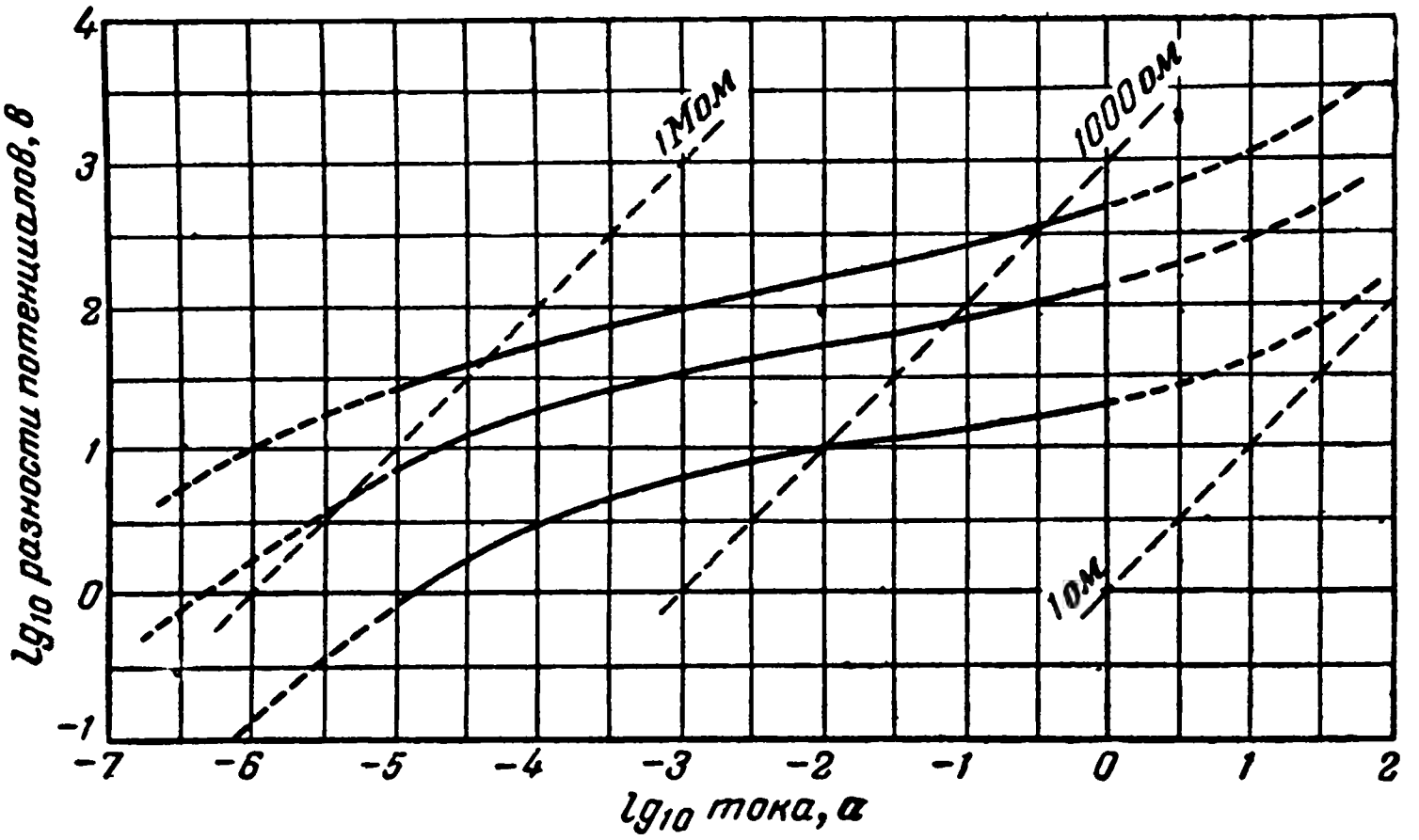


Рис. 5-21. Характеристики, показывающие изменение свойств нелинейного элемента из карбида кремния диаметром 19 *мм* и толщиной 1 *мм* при изменении режима термообработки.

ластей дает выпрямляющие контакты. Большое количество соединенных в беспорядке последовательных и параллельных цепей из выпрямительных элементов приводит к образованию сложной характеристики, подобной показанной на рис. 5-21. Действительно, форма такой характеристики может быть почти в точности воспроизведена, правда при много меньшем напряжении, путем параллельного включения двух диодов в противоположных направлениях. Такую цепь можно рассматривать как симметричный нелинейный элемент для низких напряжений. Добавив последовательно включенные выпрямители в каждую параллельную ветвь, можно получить более высоковольтную составную характеристику. Однако при этом вскоре стоимость и объем такого устройства станут чрезмерными и предпочтительнее окажется элемент из карбида кремния.

Эквивалентную схему для симметричного нелинейного элемента можно принять такую же, как и приведенная на рис. 5-4 для диода, но считать, что R не зависит от полярности и уменьшается с ростом напряжения при любой полярности. В пределах значительной области напряжений и токов характеристику можно приближенно выразить уравнением $I=AU^n$, где n — показатель степени — имеет величину от 3 до 5; A — постоянная, зависящая от геометрии и технологии изготовления элемента. При высоких напряжениях, возникающих, например, при переходных режимах, R уменьшается и может начать играть роль предельное сопротивление r.

ЗАДАЧИ

1. Какое должно быть удельное сопротивление в пределах запирающего слоя толщиной 10^{-5} *см*, чтобы обратное сопротивление при 1 *в* было таким же, как у выпрямителя на рис. 5-3? Площадь выпрямителя 0,16 *см*^{2}.

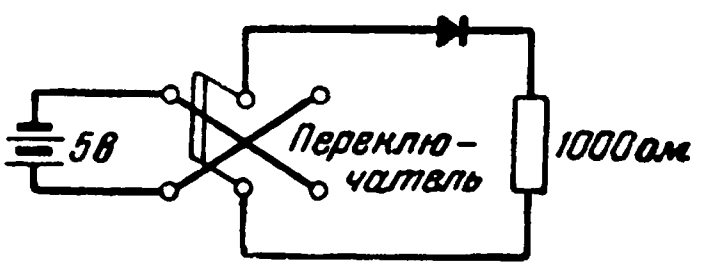


Рис. 5-22. Схема к задаче 2.

2. Каково отношение токов, проходящих при двух положениях переключателя в цепи, изображенной на рис 5-22. Диод имеет характеристики, приведенные на рис. 5-2 и 5-3. (Ответ: 45 : 1). Почему эта величина отличается от «коэффициента выпрямления при 5 *в*»?

3. Постройте характеристику выпрямителя на рис. 5-3, если элемент имеет проводимость 100 *мкмо*.

ЛИТЕРАТУРА

1. Brattain W. H, The Copper Oxide Varistor, Bell Laboratories Record, 1940, vol. 19, p. 153.

2. Henisch H. K, Metal Rectifiers, Oxford—Clarendon, 1949.

3. VanGeel Ch., The Copper Oxide Rectifier, Physica, 1934, vol. 1, p. 531; On the Cuprous Oxide Rectifier, Physica, 1934, vol. 1, p. 1143.

4. Smith I. R., Characteristics of Copper Oxide Rectifiers, Westinghouse Engineer, 1943, vol. 3, p. 85.

5. Ramsey G., The Selenium Rectifier, Electrical Engineering, 1944, vol. 63, p. 425.

6. Yarmack J. E., Selenium Rectifiers and Their Design, Electrical Communication, 1941—1942, vol. 20, p. 275.

7. Pietenpol W. J., p—n Junction Rectifier and Photocell, 1951, vol. 82, p. 121.
8. Goucher F. S, Theory and Experiment for a Germanium p—n Junction, Physical Review, 1951, vol. 81, p. 637.
9. Mott N. F., The Theory of Crystal Rectifiers, Proceedings of the Royal Society (London), 1939, vol. 171, p. 27.
10 Davydov G, On the Contact Resistance of Semiconductors, Journal of Physics USSR, 1939, vol. 1, p. 167.
11. Schottky W., Vereinfachte und erweiterte Theorie der Randschichtgleichrichter, Zeitschrift für Physik, 1941—1942, Bd 118, S. 539.
12. Prince M. B., Diffused p—n Junction Rectifiers, Bell System Technical Journal, 1956, vol. 35, p. 661.
13. McEachron K. B, Thyrite, a New Material for Lightning Arrestors, Transactions of the American Institute of Electrical Engineers, 1930, vol. 49, p. 410.
14. Grisdale R. O. Silicon Carbide Varistors, Bell Laboratories Record, 1940, vol. 19, p. 46.
15. Frosch C. J, Improved Silicon Carbide Varistors, Bell Laboratories Record, 1954, vol. 32, p. 336.
16. Dienel H. F., Silicon Carbide Varistors: Properties and Construction, Bell Laboratories Record, 1956, vol. 34, p. 407.
17. Пужай А. Н., Германиевые диоды, «Автоматика и телемеханика», 1956, 17, 2, 140—147.

ГЛАВА ШЕСТАЯ

СВОЙСТВА ВАРИСТОРОВ; ПОВЕДЕНИЕ ИХ ПРИ ВЫСОКОЙ ЧАСТОТЕ; ТОЧЕЧНО-КОНТАКТНЫЕ ПРИБОРЫ

В этой главе мы рассмотрим некоторые основные параметры, используемые при описании свойств варисторов. К ним относятся активное сопротивление, полное сопротивление, емкость и зависимость этих величин от постоянного смещения и от частоты переменного тока. Мы рассмотрим частотные ограничения, накладываемые барьерной емкостью выпрямителя на работу плоскостных диодов, и затем покажем, как эти ограничения могут быть смягчены при переходе к структурам с точечным контактом.

6-1. СОПРОТИВЛЕНИЕ ПОСТОЯННОМУ И ПЕРЕМЕННОМУ ТОКУ

При дальнейшем рассмотрении характеристик варисторов следует уточнить, что подразумевается под термином «активное сопротивление» применительно к нели-

нейному схемному элементу. Рассмотрим левую часть электрической характеристики на рис. 6-1. В любой точке характеристики, например p, можно легко определить сопротивление как отношение напряжения к току в данной точке, т. е. как U_p/I_p. Это отношение определяется также наклоном прямой, соединяющей начало координат с точкой p. Для нелинейных устройств активное сопротивление, найденное таким образом, различно при разных величинах напряжения и тока; для несимметричных нелинейных элементов оно, кроме того, различно при по-

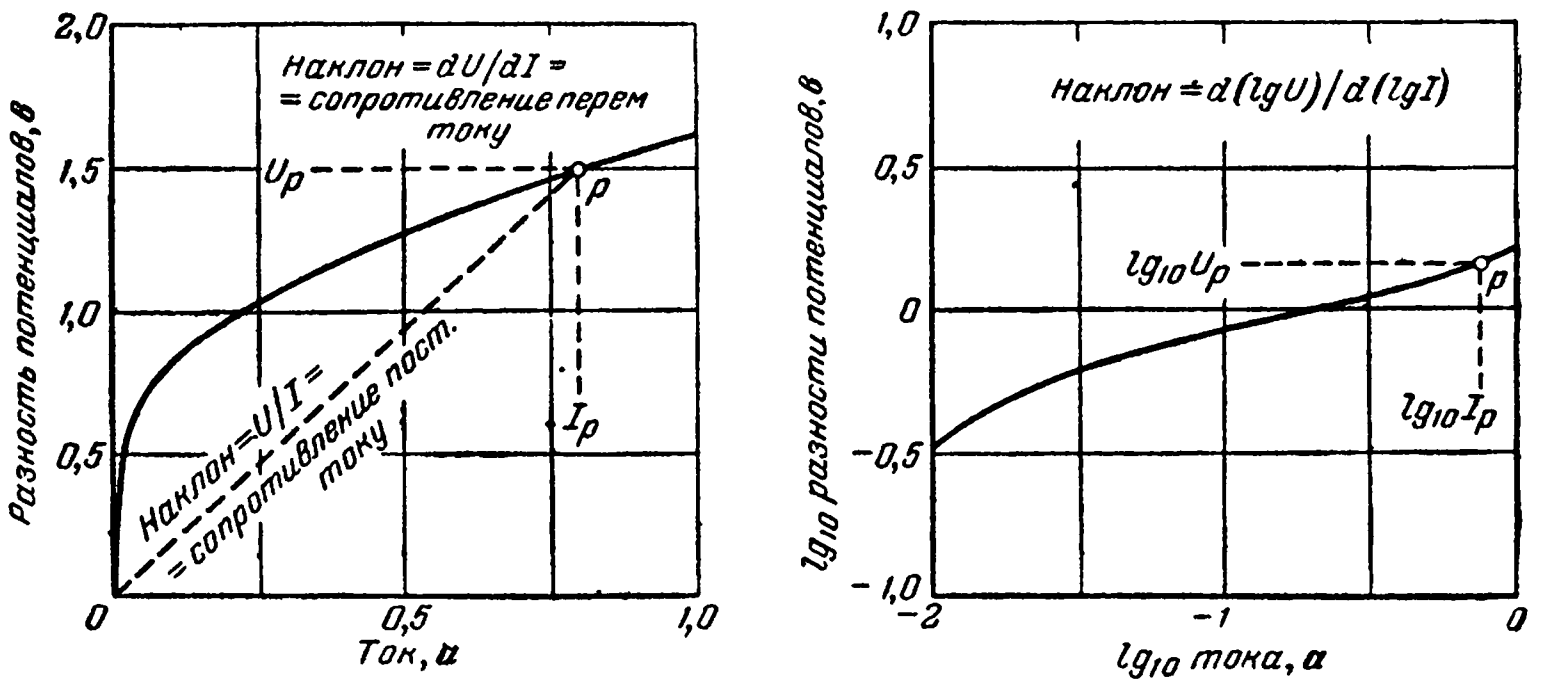


Рис. 6-1. Примеры графического определения сопротивлений постоянному и переменному токам непосредственно по характеристикам прибора.

ложительном и отрицательном напряжениях одной и той же величины. Принято считать отношение U_p/I_p сопротивлением постоянному току для данной пары значений тока и напряжения, соответствующих точке p.

Во многих случаях применения нелинейный элемент работает при постоянном смещении, определяющем рабочую точку p, и на это смещение накладывается небольшое переменное напряжение или ток. В этом случае для переменного сигнала сопротивление прибора будет иметь величину, равную не отношению U_p/I_p, а наклону характеристики в точке p, т. е. $[dU/dI]_p$. Это сопротивление называется сопротивлением переменному току, динамическим или дифференциальным [1].

[1] В советской терминологии принят термин «дифференциальное сопротивление», а динамическим называется сопротивление прибора при частотах, вызывающих изменение формы его характеристики по сравнению с формой статической характеристикой. *Прим. ред.*

Можно графически определить сопротивление как постоянному, так и переменному току в любой рабочей точке p, если имеется постоянная статическая характеристика прибора; для этого надо измерить отношения U/I и dU/dI в точке с заданным смещением. Подобным же образом, если имеется логарифмическая характеристика, можно определить величину сопротивления постоянному току в любой рабочей точке, непосредственно измерив величины U и I. Определение сопротивления переменному току по логарифмическим характеристикам требует, однако, знания величин U, I и наклона логарифмической характеристики, так как наклон логарифмической характеристики равен:

$$\frac{d(\lg U)}{d(\lg I)} = \frac{I}{U} \cdot \frac{dU}{dI} .$$

Отсюда сопротивление переменному току в точке P равно:

$$\left[\frac{dU}{dI}\right]_p = \left[\frac{U}{I} \cdot \frac{d(\lg U)}{d(\lg I)}\right]_p = \frac{U_p}{I_p} , \qquad (6\text{-}1)$$

умноженному на наклон логарифмической характеристики в точке P.

Правый график на рис. 6-1 поясняет определение сопротивления переменному току по логарифмической характеристике того же прибора, статическая характеристика которого дана на левом графике.

Сопротивление прибора переменному току может быть измерено экспериментально с помощью моста переменного тока, приспособленного для наложения на исследуемый прибор постоянного смещения желаемой величины. Измеряемый сигнал переменного напряжения поддерживается небольшим, обычно в несколько сотых долей вольта. На рис. 6-2 сплошной кривой показана зависимость сопротивления меднозакисного выпрямителя переменному току от напряжения постоянного смещения. Для сравнения пунктиром проведена характеристика сопротивления того же элемента постоянному току. Для варистора, наклон логарифмической характеристики которого при прямом токе никогда не превышает единицы, кривая сопротивления постоянному току для прямого направления всегда расположена выше кривой сопротивле-

ния переменному току. При напряжении смещения, равном или близком к нулю, на участке, где характеристика в прямом и обратном направлениях приближается к линейной, прибор ведет себя, как омическое сопротивление;

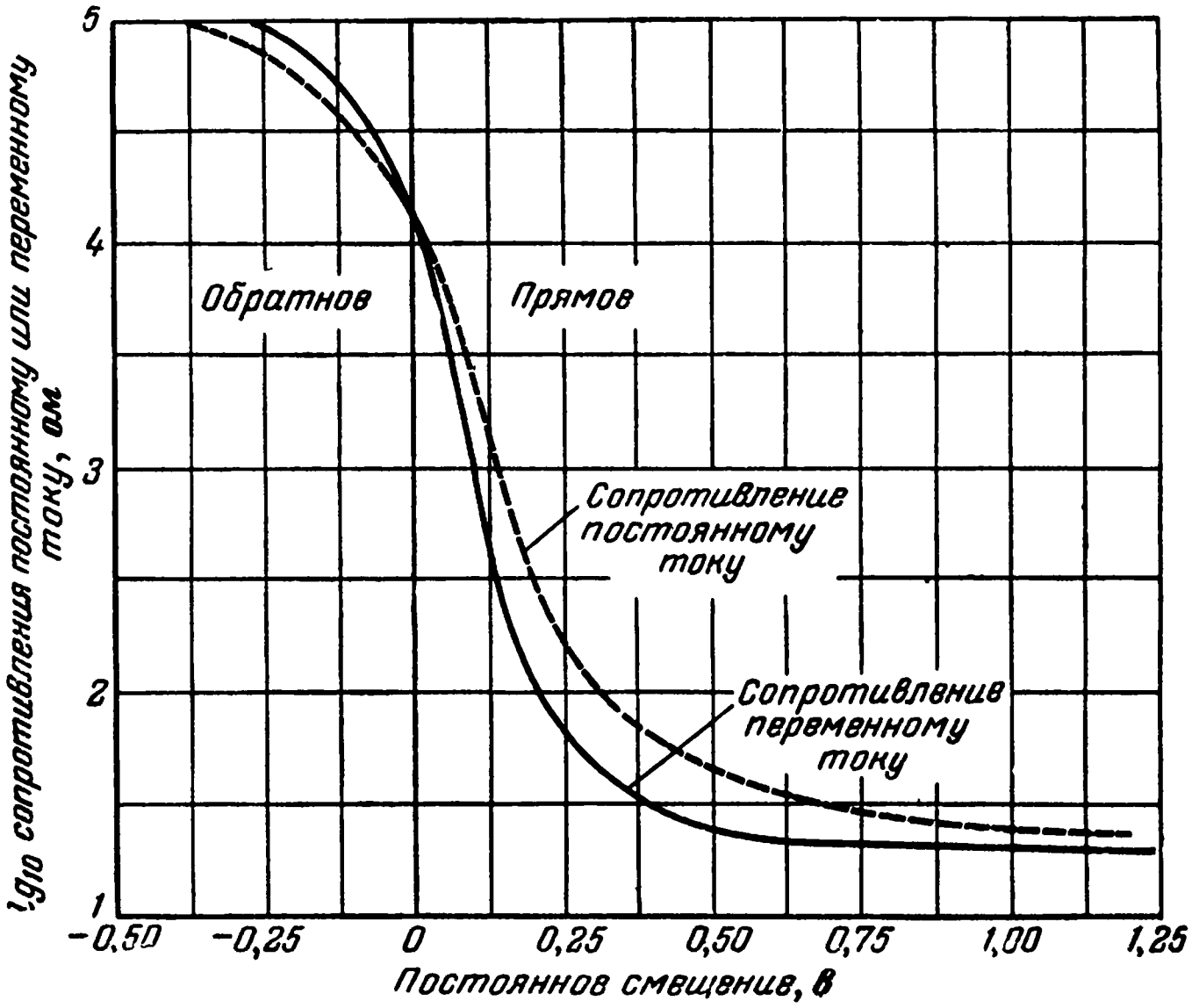


Рис 6-2 Зависимость сопротивления постоянному и переменному токам низкой частоты от напряжения постоянного смещения для меднозакисного выпрямителя диаметром ~ 5 *мм*

здесь $n=1$, и обе кривые пересекаются, как показано на графике, причем сопротивление переменному току становится больше, чем постоянному току, и сохраняет это значение в некоторых пределах значений обратного смещения. Кривые снова пересекаются при таком напряжении обратного смещения, когда обратное сопротивление постоянному току максимально.

6-2. БАРЬЕРНАЯ ЕМКОСТЬ

Ранее уже было указано, что емкость выпрямительного барьерного слоя C, так же как и его сопротивление R, зависит от величины постоянного смещения. Выражение, описывающее эту зависимость для металло-по·

лупроводникового контактного выпрямителя, дал Шоттки [Л. 1]:

$$C = \frac{A}{\sqrt{U_0 - U}} \ \phi/1 \ см^2, \qquad (6\text{-}2)$$

где A и U — константы выпрямителя и U — постоянное напряжение смещения, которое считается положительным, если оно приложено в прямом направлении, и отрицательным — в обратном. U_0 имеет размерность вольт; физически оно обозначает контактную разность потенциалов между металлом и полупроводником в выпрямительном слое. Константа A может быть вычислена теоретически:

$$A = \sqrt{\frac{\varepsilon e N}{8\pi}} \cdot 1{,}11 \cdot 10^{-12}, \qquad (6\text{-}3)$$

где ε — диэлектрическая проницаемость полупроводника; e — заряд электрона и N — концентрация примесей в области запирающего слоя, обусловливающая тип проводимости полупроводника.

Физическая модель, положенная в основу вывода уравнения (6-2), предполагает, что между обкладками емкости, образованной металлическим контактом и массой полупроводника, помещен тонкий слой диэлектрика высокого сопротивления (полупроводник в запирающем слое). Толщина запирающего слоя и, следовательно, величина емкости зависит от напряжения смещения. Толщина слоя меньше при прямых напряжениях и больше при обратных. Для прямых напряжений, величина которых приближается к U_0, величина C теоретически должна приближаться к бесконечности. На практике такое положение неосуществимо, как бы ни было велико прямое напряжение, подведенное к внешним выводам прибора, поскольку падение напряжения IR на включенном последовательно прямом сопротивлении всегда ограничивает фактическое напряжение, приходящееся на запирающий слой, величиной меньше U_0. Из уравнения (6-2) видно, что если измерить и построить график $1/C^2$ как функцию напряжения смещения U, то должна получиться прямая линия. На рис. 6-3 показан график такой функции для меднозакисного и селенового выпрямителей. Емкости из-

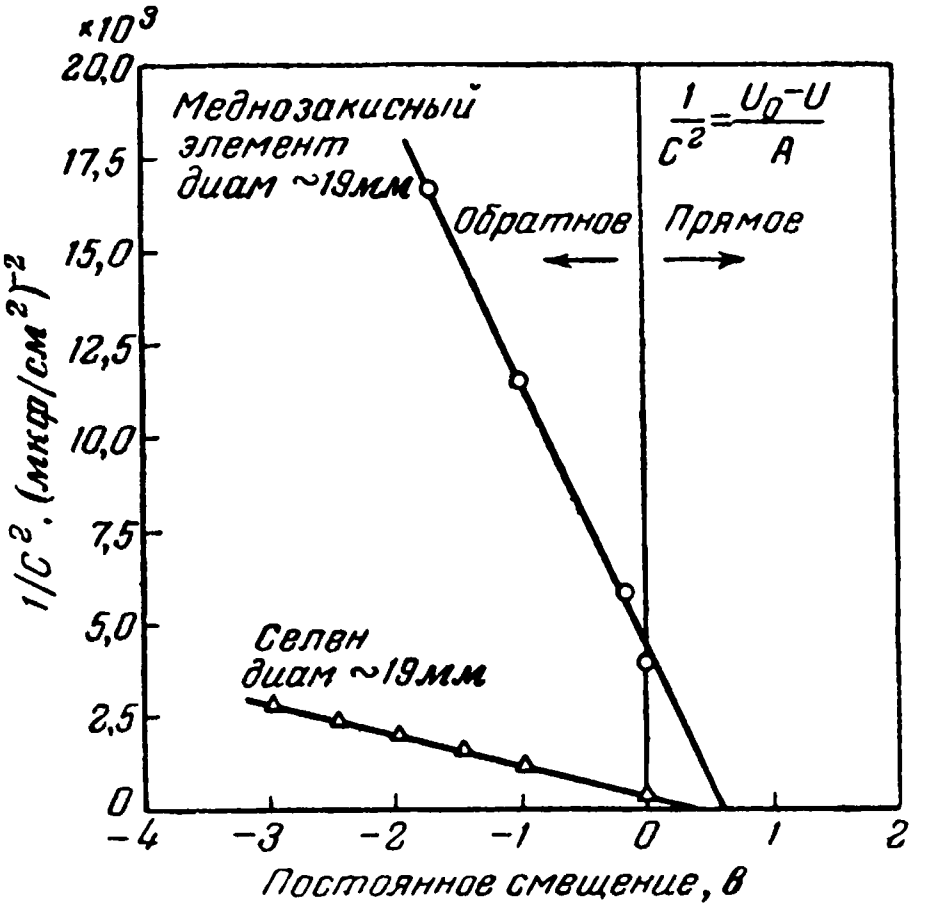


Рис. 6-3. Емкость перехода зависит от приложенного напряжения по уравнению Шоттки.

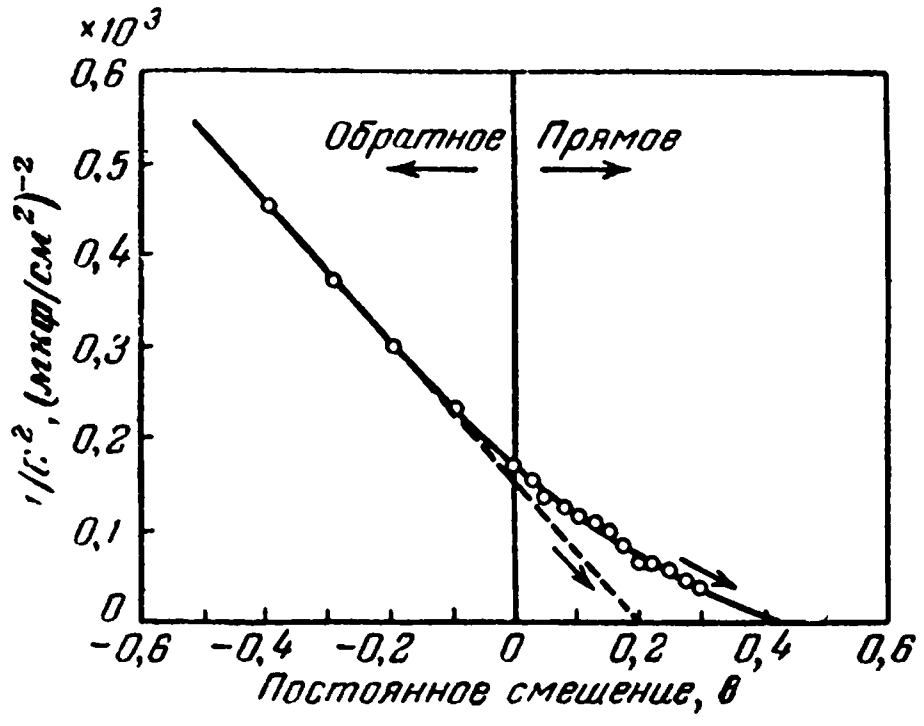


Рис 6-4 Зависимость емкости перехода от приложенного прямого напряжения не согласуется с линейным законом Шоттки для селенового выпрямителя (диаметр ~ 19 *мм*).

мерены с помощью низкочастотного моста переменного тока, в котором предусмотрена возможность подачи постоянного смещения на исследуемый прибор. Графики, построенные по экспериментально снятым точкам, хорошо согласуются с прямыми линиями, вытекающими из уравнения Шоттки. Из точек пересечения продолженных прямых с осью абсцисс можно в принципе определить величины U_0 для соответствующего прибора и из наклонов этих линий рассчитать величины A и, следовательно, N.

Для селенового выпрямителя величины емкостей, измеренные при прямом постоянном смещении, не попадают на линейное продолжение графика, снятого при обратных смещениях. На рис. 6-4 построена прямая для этого выпрямителя по уравнению Шоттки для обоих направлений. Отклонение экспериментальной кривой от прямой линии указывает, что либо величина U_0 зависит от смещения, либо величина N неодинакова в различных слоях выпрямляющего барьерного слоя. На основании ряда других признаков можно заключить, что второе предположение более вероятно.

6-3. ПОЛНОЕ СОПРОТИВЛЕНИЕ ВЫПРЯМИТЕЛЯ И ЕГО ПОВЕДЕНИЕ ПРИ МАЛЫХ ПЕРЕМЕННЫХ СИГНАЛАХ

Если выпрямитель работает при малых переменных сигналах и некотором постоянном смещении, емкость перехода C представляет собой параллельно включенное реактивное сопротивление, влияние которого сказывается при повышении частоты во все большем шунтировании сопротивления перехода и, таким образом, ухудшает нелинейные свойства прибора. Такой результат легко предсказать сразу, взглянув на эквивалентную схему выпрямителя для переменного тока (рис. 6-5). Она оказывается подобной эквивалентной схеме для постоянного тока (рис. 5-4), но отличается тем, что нелинейная составляющая R_{ac} теперь заменяется сопротивлением перехода переменному току в рабочей точке, соответствующей принятому смещению.

Чтобы выяснить природу ухудшения нелинейных свойств выпрямителя при повышении частоты, рассмотрим выражение для полного сопротивления эквивалентной схемы выпрямителя. Пренебрегая возможным влия-

нием краевой утечки, полное сопротивление можно представить в виде:

$$Z = \left[\frac{r + R_{ac} + rR_{ac}^2C^2\omega^2 - jR_{ac}^2C\omega}{R_{ac}^2C^2\omega^2 + 1}\right]_p, \qquad (6\text{-}4)$$

где ω — угловая частота (в 2π раз бóльшая числа периодов в секунду); индекс p обозначает, что все величины соответствуют постоянному смещению в данной рабочей точке.

При очень низких частотах членами, содержащими ω, в уравнении (6-4) можно пренебречь, и тогда $Z \approx r + R_{ac}$, что выражает условия, отраженные в статических

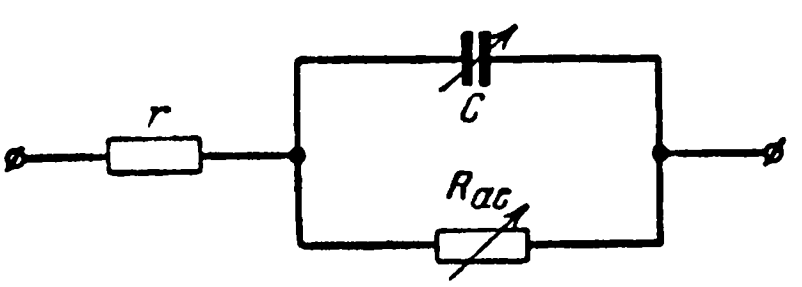


Рис. 6-5 Эквивалентная схема выпрямителя для переменного тока подобна схеме для постоянного тока, но величина сопротивления перехода переменному току $R_{пер}$ отличается от величины сопротивления постоянному току R. Стрелки обозначают, что соответствующие величины изменяются при изменении напряжения постоянного смещения

характеристиках, приведенных и рассмотренных выше. С другой стороны, при высоких частотах уравнение (6-4) можно представить в виде $Z \approx r - j/C\omega$, которое при очень высоких частотах превращается в $Z \approx r$. В последнем случае сопротивление запирающего слоя R_{ac} полностью зашунтировано и, следовательно, нелинейные свойства выпрямителя нарушены. R_{ac} не фигурирует ни в одном из уравнений для высоких частот.

Интересно исследовать характер изменения величины полного сопротивления (6-4) в зависимости от частоты, в частности в области, заключенной между рассмотренными выше крайними верхними и крайними нижними частотами. Хотя R_{ac} и C на эквивалентной схеме выпрямителя включены параллельно, при расчете цепей переменного тока принято рассматривать полное сопротивле-

ние цепи как векторную сумму последовательно включенных активного сопротивления R_s и реактивного сопротивления X_s, что дает:

$$Z = R_s + jX_s.$$

Из (6-4) следует, что

$$\left.\begin{aligned} R_s &= \left[\frac{r + R_{ac} + rR_{ac}^2C^2\omega^2}{R_{ac}^2C^2\omega^2 + 1}\right]_p; \\ X_s &= \left[\frac{R_{ac}^2C\omega}{R_{ac}^2C^2\omega^2 + 1}\right]_p. \end{aligned}\right\} \qquad (6\text{-}5)$$

Если построить зависимости R_s и X_s от частоты в логарифмическом масштабе (по обеим осям), получатся кривые, изображенные на рис. 6-6. Постоянное смещение,

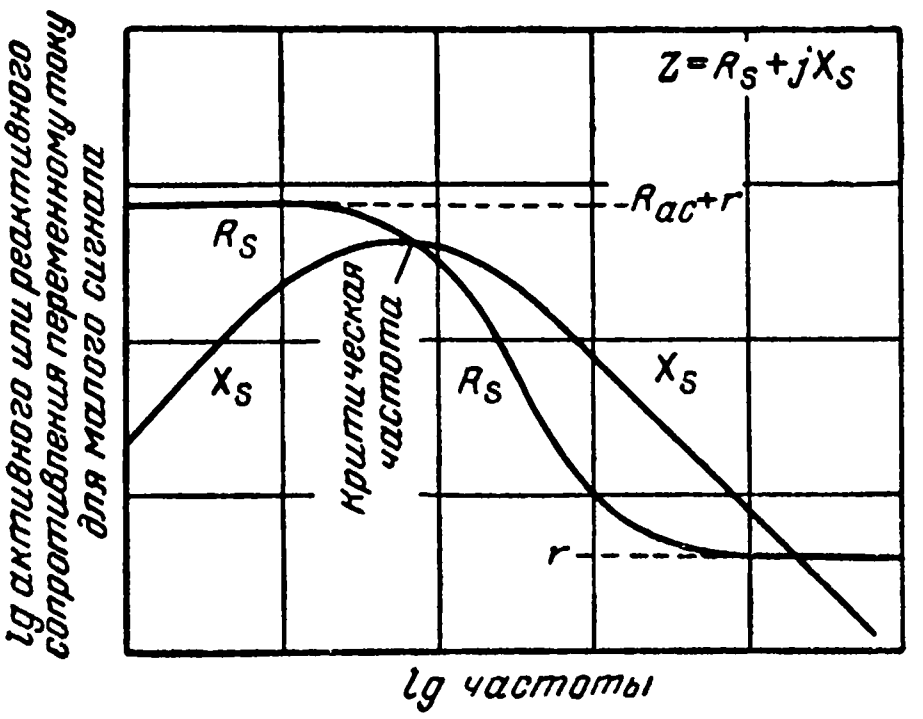


Рис 6-6. Кривые зависимости последовательных активных и реактивных сопротивлений от частоты, полученные по эквивалентной схеме выпрямителя для переменного тока

для которого построены эти характеристики, соответствует небольшому постоянному току в прямом направлении — режиму, при котором R_{ac} в 200—300 раз больше, чем r. Хотя X_s имеет отрицательный знак, на графике эта величина для сравнения представлена как положительная.

Заслуживают внимания некоторые особенности этих кривых:

1. При очень низких частотах R_s имеет величину $R_{ac}+r$ и не зависит от частоты. Мгновенная рабочая точка, соответствующая работе на низкой частоте, просто перемещается вперед и назад по статической характеристике прибора, незначительно отклоняясь от рабочей точки, соответствующей постоянному смещению.

2. При очень высоких частотах R_s равно приблизительно r и также не зависит от частоты. В этом случае

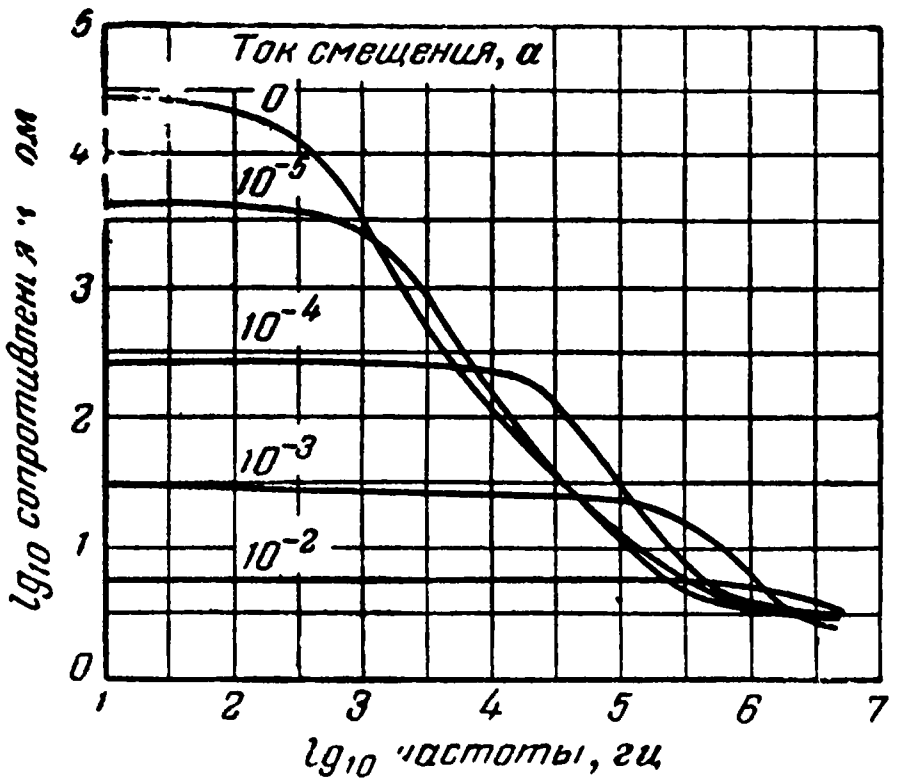


Рис 6-7 Зависимость от частоты малого сигнала последовательного сопротивления R_s для меднозакисного выпрямителя диаметром 12 *мм* Сравните эти кривые с кривыми на рис 6-6 для R_s, рассчитанными по эквивалентной схеме выпрямителя переменного тока

сопротивление перехода переменному току полностью зашунтировано реактивным сопротивлением емкости перехода и прибор ведет себя по отношению к переменному току, как активное сопротивление r.

3. При промежуточных частотах R_s уменьшается по мере повышения частоты пропорционально $1/\omega^2$. Это дает наклон кривой, равный 2 в логарифмических координатах.

4. X_s линейно возрастает при повышении частоты в области низких частот и уменьшается пропорционально $1/C\omega$ в области высоких частот, переходя через максимум при частоте $\omega=1/R_{ac}C$. В логарифмических координатах кривая симметрична относительно этой частоты и стремится к асимптотам с наклонами $+1$ и -1.

5. Кривые для R_s и X_s пересекаются друг с другом приблизительно при этой же частоте, причем как R_s, так и X_s в точке пересечения весьма близки к $R_{ac}/2$. Частота, при которой $\omega = 1/R_{ac}C$, называется критической частотой выпрямителя в данной рабочей точке (при данном постоянном смещении). Она имеет физический смысл частоты, при которой действительная часть полного сопротивления выпрямителя оказывается вдвое ниже своего зна-

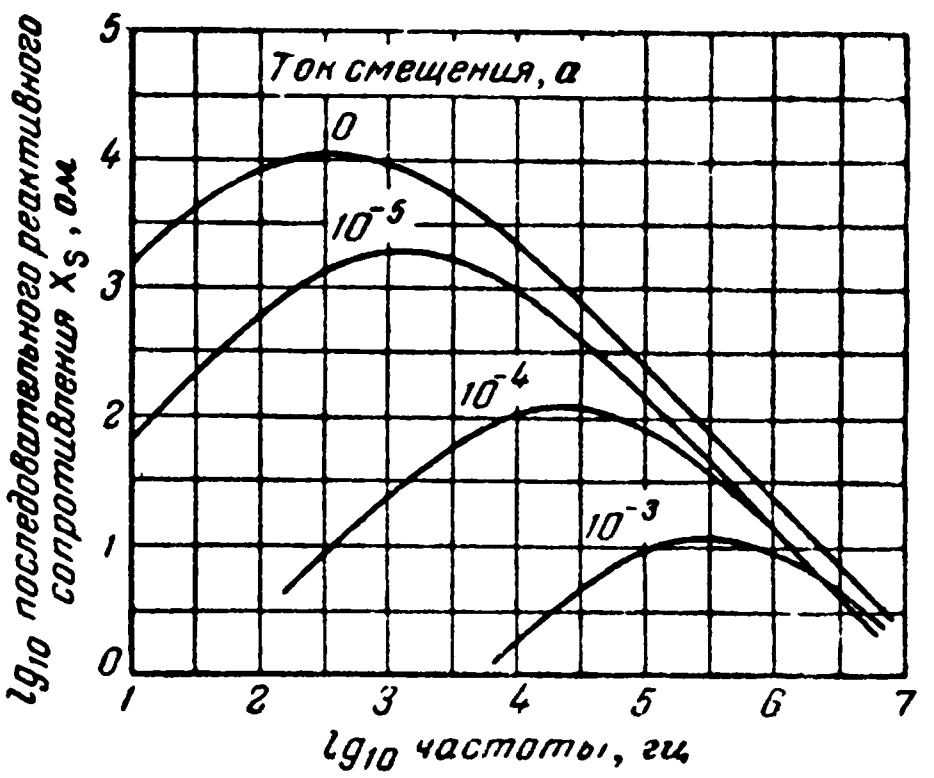


Рис. 6-8. Частотные зависимости последовательного реактивного сопротивления X_s при соответствующих постоянных токах смещения для меднозакисного выпрямителя, характеристики которого для R_s даны на рис. 6-7. Сравните с идеальными расчетными кривыми для X_s, приведенными на рис. 6-6.

чения на низких частотах, и, таким образом, указывает верхний предел, при котором частотная характеристика выпрямителя при слабых сигналах переменного тока может считаться приблизительно горизонтальной.

Качественные особенности кривых на рис. 6-6 проявляются в измеренных частотных характеристиках типичных выпрямителей. Так, например, на рис. 6-7 и 6-8 показаны экспериментально снятые кривые для R_s и X_s как функции частоты для меднозакисного выпрямителя диаметром 12,5 *мм* при нескольких значениях тока постоянного смещения [Л. 2]. Из сравнения соответствующих кривых для R_s и X_s видно хорошее совпадение с теорети-

чески вычисленными значениями для простой эквивалентной схемы. Однако отклонения некоторых кривых для R_s от теоретических кривых на рис. 6-6 показывают, что хотя рассмотренная нами простая эквивалентная схема в основном правильна и пригодна для большинства инженерных целей, она все же не является исчерпывающе полной во всех деталях.

Для плоскостных выпрямителей критическая частота теоретически не зависит от площади выпрямителя, так как при увеличении площади R и r уменьшаются, а C увеличивается в одинаковое число раз. Это предположение в основном выполняется, и наблюдаемые отклонения можно хорошо объяснить. Иногда площадь выпрямителя так мала, что величину r приходится рассматривать отчасти как распределенное сопротивление (см. § 6-5). Кроме того, в особенности для меднозакисных выпрямителей, в эквивалентную схему приходится добавлять краевую утечку, которая не принималась в расчет в нашем теоретическом анализе.

6-4. ПОВЕДЕНИЕ ВЫПРЯМИТЕЛЕЙ ПРИ БОЛЬШИХ ПЕРЕМЕННЫХ СИГНАЛАХ

При использовании выпрямителя для преобразования переменного тока в постоянный и в большинстве случаев его работы как модулятора несущей частоты сигнал имеет достаточно большую амплитуду для перемещения рабочей точки выпрямителя в широких пределах прямых и обратных напряжений и токов. Анализ работы выпрямителя при малых переменных сигналах, проведенный выше, оказывается неприменимым для больших сигналов, и мы исследуем свойства выпрямителей при больших переменных сигналах другими методами.

Наивысшая частота, при которой выпрямитель, действующий с большим переменным сигналом, может быть эффективно использован, как и в случае малых сигналов, ограничена шунтирующим действием емкости перехода. В результате при повышении рабочей частоты постепенно ухудшается коэффициент выпрямления прибора. Чтобы показать влияние этого шунтирования на форму характеристики и коэффициент выпрямления, на рис. 6-9 представлен ряд характеристик выпрямигеля, соответствующих постепенно повышающимся частотам. Строго говоря, эти кривые не являются теоретически

правильными, так как они представляют как действительную, так и мнимую части полного сопротивления на одной и той же координатной плоскости ($\lg U$—$\lg I$), но могут служить графической иллюстрацией при изучении рассматриваемого вопроса. Если рабочая частота достаточно высока, емкостное сопротивление запирающего слоя оказывается малым и шунтирует переход на участках характеристики, соответствующих высокому сопротивлению. Результирующая характеристика выпрямителя при этом сдвигается от кривой *I* к кривой *II*.

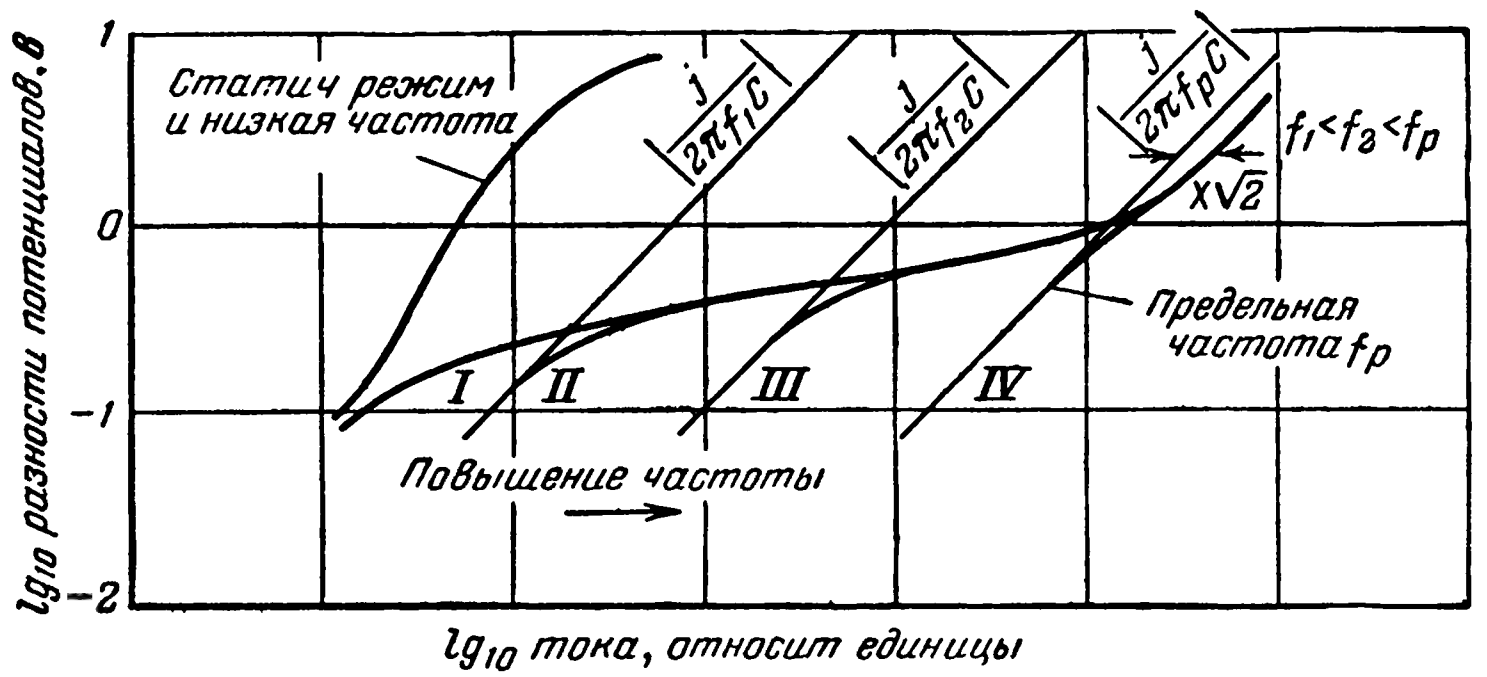


Рис. 6-9 Понижение эффективного коэффициента выпрямления по мере повышения рабочей частоты

На участке обратных токов характеристики и в области малых напряжений на ветви прямых токов основную роль играет реактивное сопротивление запирающего слоя, которое становится все меньшим и меньшим по мере повышения частоты (кривая *III*). В то же время эффективный коэффициент выпрямления снижается. При некоторой частоте f_p эффективный коэффициент выпрямления падает до $\sqrt{2}$ (кривая *IV*) и прибор практически теряет свои качества как нелинейный элемент. Рассмотрение и несложный анализ эквивалентной схемы показывают, что если C или r уменьшить, то предельная частота f_p повысится.

Чтобы получить количественную картину этого влияния частоты [Л. 2], пренебрежем сопротивлением краевой утечки в эквивалентной схеме выпрямителя и предположим, что на барьерное сопротивление наложено переменное напряжение низкой частоты, изменяющее барьерное

сопротивление между величинами $R_{обр}$ и $R_{пр}$. Коэффициент выпрямления будет тогда выражаться так:

$$\textit{Коэффициент выпрямления} = \frac{R_{обр} + r}{R_{пр} + r}.$$

Если это напряжение достаточно велико, чтобы рабочая точка заходила далеко вверх по идущей под углом 45° предельной линии прямого тока, то сопротивлением $R_{пр}$ можно пренебречь по сравнению с r; тогда

$$\textit{Коэффициент выпрямления} \approx \frac{R_{обр} + r}{r}.$$

Пусть частота повышается до тех пор, пока реактивное сопротивление емкости запирающего слоя не станет много меньше $R_{обр}$. Тогда это реактивное сопротивление будет определять характеристику выпрямителя при обратном токе и вместо $R_{обр}$ в последнем уравнении нужно подставить реактивное сопротивление $-j/\pi fC$, в результате чего получим:

$$\textit{Коэффициент выпрямления} \approx \frac{-\frac{j}{2\pi fC} + r}{r}.$$

Теперь определим пороговую частоту f_p *, при которой реактивное сопротивление емкости перехода становится равным предельному сопротивлению тела выпрямителя, т. е.

$$\left| -\frac{j}{2\pi f_p C} \right| = r. \qquad (6\text{-}6)$$

При этой предельной частоте коэффициент выпрямления принимает величину

$$\textit{Коэффициент выпрямления} \approx \frac{-jr + r}{r} \approx \sqrt{2}. \qquad (6\text{-}7)$$

Предельная частота является удобным критерием качества выпрямителя, поскольку она определяет верхний предел диапазона частот, в котором данный полупровод-

---

* Эту предельную частоту не следует смешивать с критической частотой переменного сигнала малой амплитуды, которая рассмотрена в последнем параграфе.

никовый прибор может эффективно работать как выпрямитель или модулятор сигналов большой амплитуды. Расчет предельной частоты по уравнению (6-6) не является простым и точным, поскольку величина C зависит от напряжения и, следовательно, принимает различные значения в течение каждого периода переменного напряжения. Значения предельных частот могут быть определены экспериментально. Они колеблются от 10 до 100 *кгц* для селеновых выпрямителей и от 1 до 2 *Мгц* для меднозакисных выпрямителей. Последние приборы успешно используются в качестве модуляторов для групповых каналов систем уплотнения телефонных линий связи на частотах до 2 *Мгц*, но оказываются очень малоэффективными в качестве силовых выпрямителей на этих частотах.

6-5. ВЫПРЯМИТЕЛИ С РАСПРЕДЕЛЕННЫМ СОПРОТИВЛЕНИЕМ И ТОЧЕЧНО-КОНТАКТНЫЕ

Имеется много случаев применения, когда плоскостные выпрямители непригодны из-за того, что их предельные частоты недостаточно высоки. Поэтому естественно поставить вопрос о том, как можно повысить предельную частоту выпрямителя. Рассмотрение уравнения (6-6) показывает, что f_p пропорционально $1/rC$. Если r либо C или обе величины вместе можно изменить так, чтобы уменьшить произведение rC, то в результате будет достигнута более высокая предельная частота. Для плоскостных выпрямителей, в которых использован определенный полупроводник с заданными удельным сопротивлением и толщиной слоя, произведение rC нельзя уменьшить путем изменения площади выпрямительного контакта, так как величины r и C, завися от размеров площади, всегда изменяются при изменении последней в противоположных направлениях.

Ограничения рабочей частоты, обусловленные произведением rC, могут быть ослаблены приданием контакту точечной формы вместо плоской, которую мы до сих пор рассматривали. К типу точечных относятся выпрямители, размеры выпрямительной контактной поверхности которых малы сравнительно с толщиной слоя полупроводника. Такой выпрямитель осуществляется обычно посредством прижатия к поверхности полупроводника проволочного острия, как показано на рис. 6-10.

Изготовление точечных выпрямителей мы рассмотрим несколько ниже, а сейчас укажем, какие замечательные особенности точечных выпрямителей позволяют применять их на таких частотах, при которых плоскостные приборы оказываются бесполезными.

Конфигурация линий тока в полупроводнике вблизи выпрямительного контакта различна у плоскостных и точечных выпрямителей, как поясняют два эскиза на рис. 6-11. В плоскостных выпрямителях линии тока параллельны и проходят прямолинейно через слой полупроводника от выпрямительного контакта к противолежащему электроду, причем эквипотенциальные поверхности представляют собой семейство параллельных плоскостей. Наоборот, в точечном выпрямителе линии тока направлены приблизительно по радиусам в толще полупроводника вблизи точечного контакта, что соответствует примерно полусферическим эквипотенциальным поверхно-

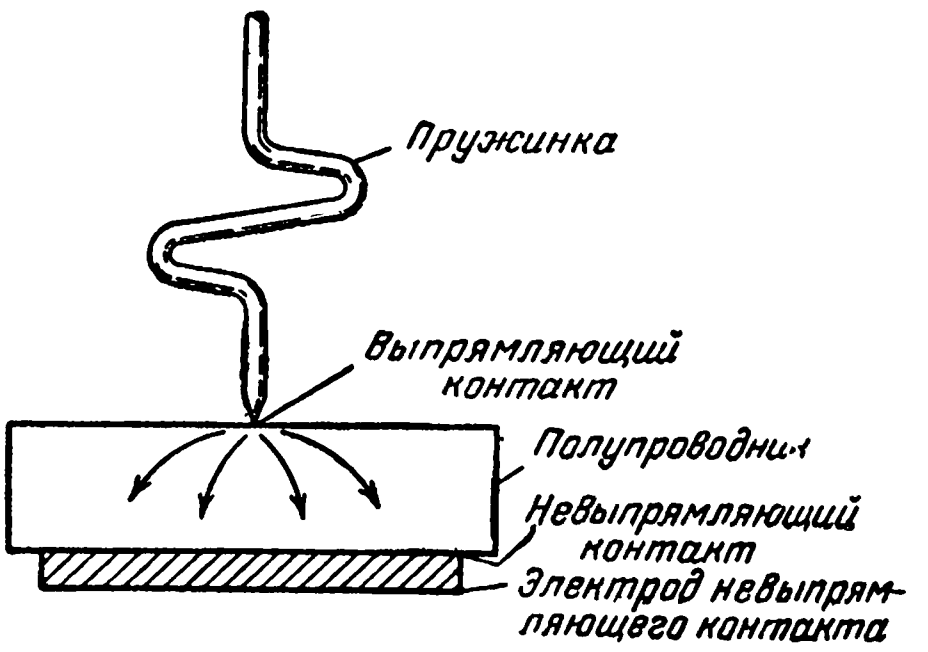


Рис. 6-10. Основные детали точечного выпрямителя.

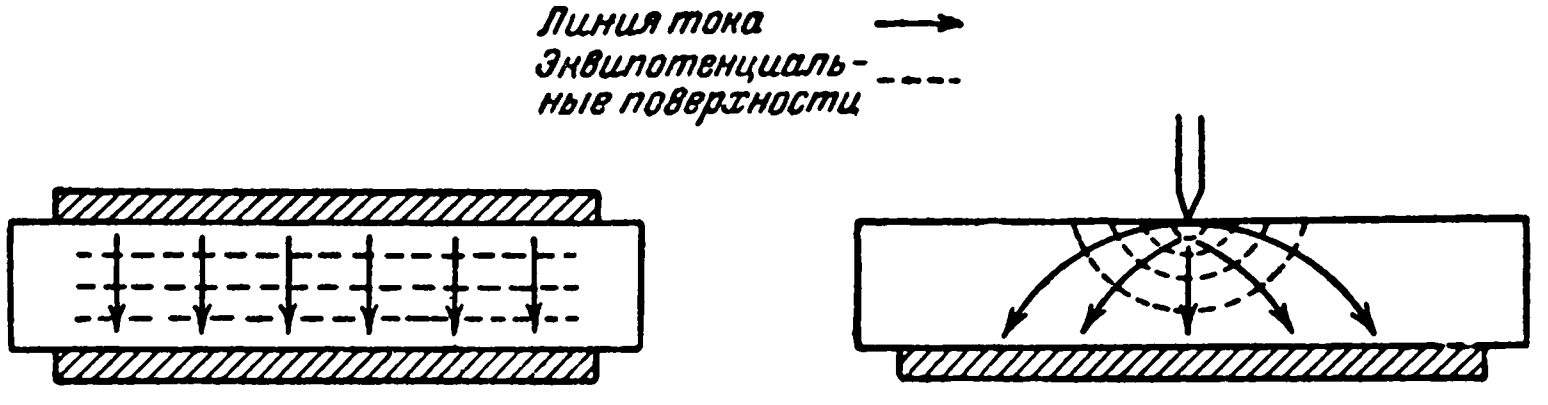


Рис. 6-11. В плоскостном выпрямителе (слева) линии тока практически параллельны; в точечном выпрямителе (справа) линии тока почти радиальны. Точечная геометрия обладает преимуществами, обусловленными «эффектом распределенного сопротивления».

стям. Это радиальное распределение тока приводит к тому, что предельное сопротивление r массы полупроводника для прямого тока не зависит от размера или формы полупроводникового диска, а почти целиком определяется свойствами малого объемного элемента полупроводника, непосредственно примыкающего к точечному контакту и имеющего радиус, равный двум-трем диаметрам точечного контакта. Можно показать аналитически [Л. 3], что сопротивление дискообразного контакта диаметром d с полубесконечным твердым телом (ограниченным плоскостью), имеющим удельное сопротивление ρ, равно:

$$r = \frac{\rho}{2d} . \qquad (6\text{-}8)$$

Таким образом, если постепенно уменьшать размеры контакта, то предельное сопротивление r выпрямительного элемента для прямого тока будет увеличиваться обратно пропорционально *радиусу* контакта, а не его *площади*. Это явление носит название «эффекта распределенного сопротивления». Однако при этом сопротивление R выпрямителя для обратного тока, природа которого заложена в существовании и свойствах перехода, должает увеличиваться обратно пропорционально *площади* контакта. Эффект распределенного сопротивления становится заметным, если диаметр контакта меньше примерно половины толщины слоя полупроводника. Если диаметр контакта меньше примерно одной десятой толщины слоя полупроводника, эффект распределенного сопротивления проявляется с полной силой [1] и уравнение (6-8) становится вполне точным.

Влияние распределенного сопротивления на форму характеристик выпрямителя показано на рис. 6-12. Справа изображена логарифмическая характеристика A для плоскостного круглого прижимного контакта с полупроводником. Эта характеристика соответствует сравнительно плохому коэффициенту выпрямления, что характерно

---

[1] Другим интересным следствием эффекта распределенного сопротивления является то, что сопротивление обратной цепи (через землю) между двумя станциями проводной связи определяется главным образом размерами заземляющих электродов и влажностью грунта вблизи них, оставаясь почти неизменным при расстояниях между станциями, скажем, 100 *м* или 1 000 *км*.

для прижимных контактов большой площади. Если теперь этот контакт уменьшить до одной десятой его прежнего диаметра и при этом сохранятся соотношения размеров, характерные для плоскостного контакта, как это было определено выше, то новая характеристика *Б* получится простым сдвигом первоначальной характеристики в сторону меньших токов на расстояние, соответствующее коэффициенту 100. Теперь пусть контакт будет

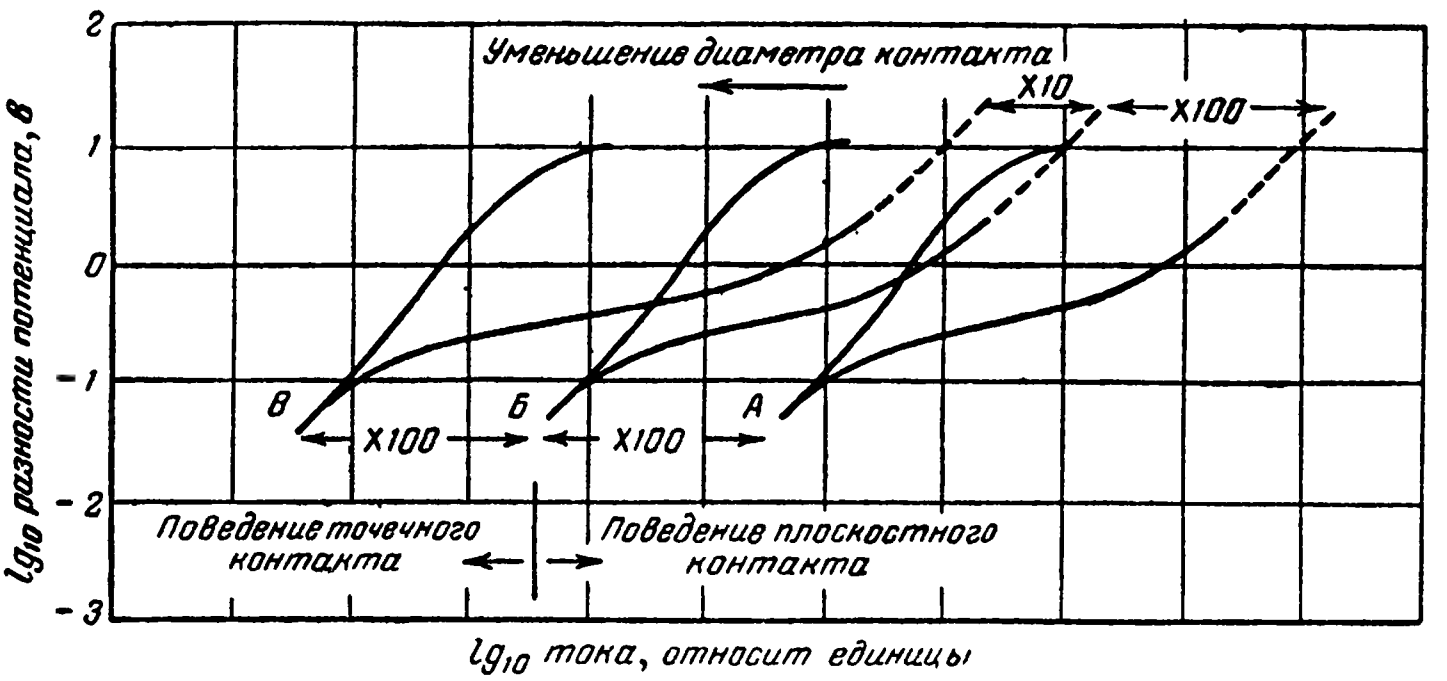


Рис 6-12. Превращение характеристик плоскостного диода в характеристики точечного диода по мере уменьшения диаметра контакта: коэффициент выпрямления улучшается, повышается предельная частота.

уменьшен еще в 10 раз и предположим, что при этом размеры контакта станут такими, что будет сказываться эффект распределенного сопротивления. Соответствующая характеристика *В* также сместится дальше на расстояние, соответствующее уменьшению токов в 100 раз, за исключением области прямого предельного сопротивления, которая сместится на величину, соответствующую уменьшению тока лишь в 10 раз, как это следует из уравнения (6-8). Следовательно, характеристика *В* соответствует примерно в 10 раз большему коэффициенту выпрямления при напряжениях выше 1 *в*, чем характеристики *А* или *Б*. При дальнейшем уменьшении диаметра контакта будет иметь место дальнейшее прямо пропорциональное ему смещение участка характеристики, соответствующего прямому предельному сопротивлению, тогда как остальная часть характеристики смещается к токам, уменьшенным пропорционально квадрату изменения диаметра контакта. Таким образом, коэффициент

выпрямления будет все более возрастать. Применение точечных контактов позволяет, таким образом, получить более высокие значения коэффициента выпрямления, чем у плоскостных прижимных контактов. Разумеется, при уменьшении диаметра контакта максимальный допустимый ток выпрямителя уменьшается. Следовательно, точечные выпрямители применимы в тех случаях, когда можно пожертвовать величиной тока в пользу более высокого коэффициента выпрямления и, как мы сейчас убедимся, лучшей частотной характеристики.

Пример, приведенный на рис. 6-12, идеализирован в том отношении, что переход от характеристик, соответствующих плоскостным приборам, к характеристикам приборов с распределенным сопротивлением не происходит внезапно при некотором определенном диаметре контакта. Этот пример служит только иллюстрацией разницы зависимости характеристик от диаметра контакта для случаев плоскостной и точечной геометрии.

Рассмотрим теперь, каким образом эффект распределенного сопротивления позволяет получить более высокие величины предельных частот, чем у плоскостных выпрямителей. Величина предельной частоты обратно пропорциональна rC. Для точечного выпрямителя r обратно пропорционально *диаметру* контакта, тогда как C, будучи свойством контактного перехода, изменяется пропорционально *площади* контакта.

Таким образом,

$$f_p \sim \frac{1}{rC} \sim \frac{1}{\left(\frac{1}{d}\right)d^2} \sim \frac{1}{d} . \qquad (6\text{-}9)$$

Из этого уравнения видно, что для выпрямительных контактов достаточно малых, чтобы в них мог проявиться эффект распределенного сопротивления, предельная частота обратно пропорциональна диаметру контакта. Предельная частота германиевых точечных диодов меняется от 100 до 1 000 *Мгц*, доходя у кремниевых диодов до 10 000 *Мгц*.

На заре радиотехники в качестве полупроводников для кристаллических детекторов применялись главным образом свинцовый блеск (галенит) и серный колчедан (пирит). В ранних конструкциях этих приборов полупроводник, обычно в виде осколка неправильной формы объ-

емом в несколько кубических миллиметров, впаивался в держатель, а заостренная металлическая проволочка (пружинка) прикреплялась к подвижному рычажку, так что контакт пружинки с открытой поверхностью полупроводника осуществлялся и поддерживался за счет ее упругости при установке от руки. Эти первые детекторы

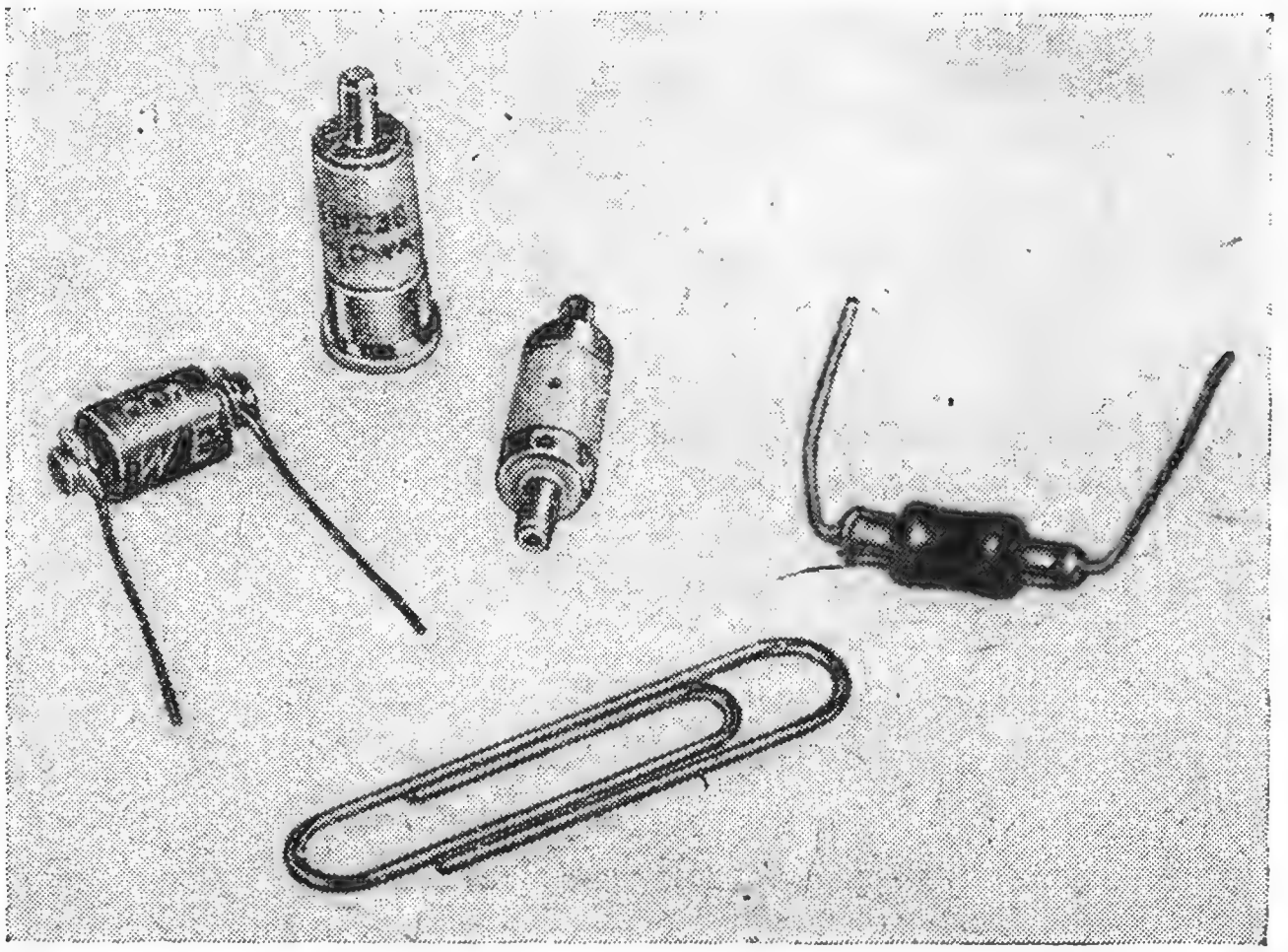


Рис 6-13 Варианты внешнего оформления точечных кристаллических диодов Слева направо — корпуса из фенолпласта, керамики, коаксиальный керамический и стеклянный запаянный баллон

имели весьма нестабильные характеристики; часто их установка полностью сбивалась, и тогда нужно было искать новую точку контакта на кристалле. Это вызывало подчас необходимость долго и кропотливо искать эту точку, так как не все они равно пригодны для детектирования. Современные точечные детекторы представляют собой значительный шаг вперед по сравнению с этими ранними приборами. В качестве полупроводников в них используется почти исключительно кремний или германий. Технологический процесс приготовления полупроводника обеспечивает образование поверхности, вся площадь которой почти одинаково пригодна для получения выпрямляющего контакта. Настроенный контакт надежно

фиксируется во время сборки и остается в таком положении все время службы прибора.

Полупроводниковые диоды в настоящее время изготовляются в нескольких конструктивных оформлениях: в керамических капсулах, в коаксиальных капсулах, в капсулах с проволочными выводами и в запаянных стеклянных баллончиках. На рис. 6-13 приведена фотография диодов этих типов. Во всех этих конструкциях геометрия полупроводника и металлического электрода в основном одна и та же. Полупроводник представляет собой прямоугольную пластинку кремния или германия, выпиленную из заготовки. Невыпрямляющий контакт с пластинкой осуществляется путем покрытия ее задней стороны соответствующим металлом с последующей припайкой к металлической детали, служащей одним из внешних электродов готового прибора. Контактная пружинка, обычно из фосфористой бронзы или вольфрама, припаивается или приваривается к держателю и механически изгибается нужным образом; конец пружинки затачивается по требуемой форме механическим или электролитическим способом. Затем при окончательной сборке изделия острие пружинки приводится в соприкосновение с подготовленной поверхностью полупроводника. На рис. 6-14 показан продольный разрез кремниевого диода в керамическом корпусе, иллюстрирующий указанные конструктивные особенности.

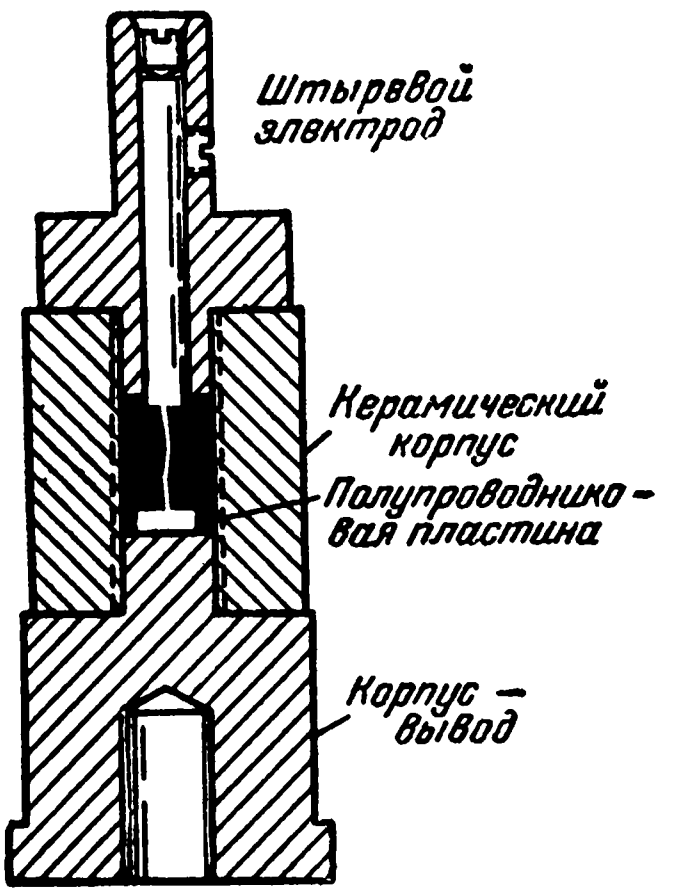


Рис 6-14 Продольный разрез полупроводникового выпрямителя.

Механическая стабильность в этом диоде достигается некоторым сжатием контактной пружинки при сборке. Улучшенная устойчивость к механическим ударам и защита выпрямительного контакта от влаги и химически активных газов достигаются заполнением полости корпуса составом, не расплавляющимся при высоких температурах и не растрескивающимся от хрупкости при низких.

6-5-1. Выпрямители со сплавным контактом

Новым методом изготовления выпрямительных контактов, приводящим к удивительно своеобразным характеристикам диодов, является сплавление или сварка [Л. 4].

В настоящее время бóльшая часть выпрямителей со сплавным переходом выполняется путем приварки конца алюминиевой проволочки к поверхности кремния [Л. 5]. При этой операции пластина кремния *n*-типа площадью примерно 1 *мм*^{2} и толщиной 0,5 *мм* помещается на электрически нагреваемую танталовую ленточку. Затем заостренный конец алюминиевой проволочки приводится в соприкосновение с верхней поверхностью пластины и вся эта система нагревается в атмосфере инертного газа. При температуре плавления эвтектики кремния и алюминия в точке контакта образуется крошечная капля расплавленного кремне-алюминиевого сплава. После этого температуру снижают, в результате чего сплав кристаллизуется. В процессе кристаллизации бóльшая часть алюминия выделяется из застывающего сплава, но в нем все же остается достаточно алюминия, чтобы придать ему проводимость *p*-типа вместо *n*-типа. В результате этого процесса алюминиевая проволочка оказывается вплавленной в полусферический островок кремния *p*-типа, включенный в массу исходного кремния *n*-типа. Полученный *p-n* переход в виде полусферической поверхности, имеющей радиус несколько сотых долей миллиметра, и является выпрямительным переходом. При этой же тепловой операции к одной из свободных поверхностей пластины припаивается золотой невыпрямляющий контакт. Затем весь элемент заключается в оболочку для механической защиты. На рис. 6-15 показан эскиз одно-

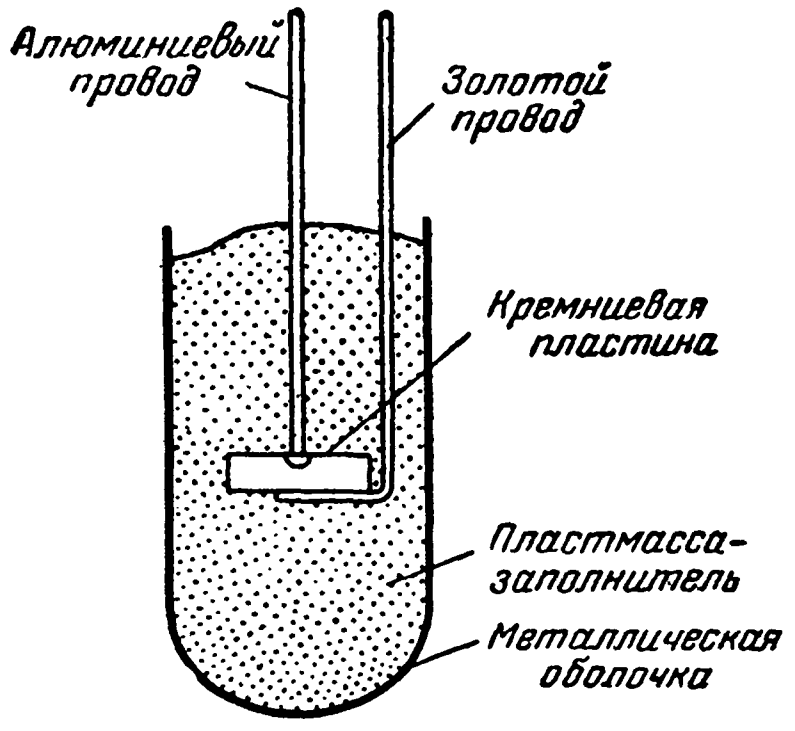


Рис. 6-15. Кремниево-алюминиевый сплавной диод, изготовляемый вплавлением кончика алюминиевой проволочки в поверхность кремниевой пластинки.

го из вариантов конструктивного оформления сплавных диодов.

На рис. 6-16 представлена характеристика такого диода. Отличительной особенностью этой характеристики является огромная величина полного сопротивления в области обратного тока насыщения, достигающая тысяч мегом. При напряжении 1 *в* достигается коэффициент выпрямления от 10^7 до 10^8. Как и для диодов с выращенным переходом, напряжение обратного пробоя может при изготовлении регулироваться в широких пределах.

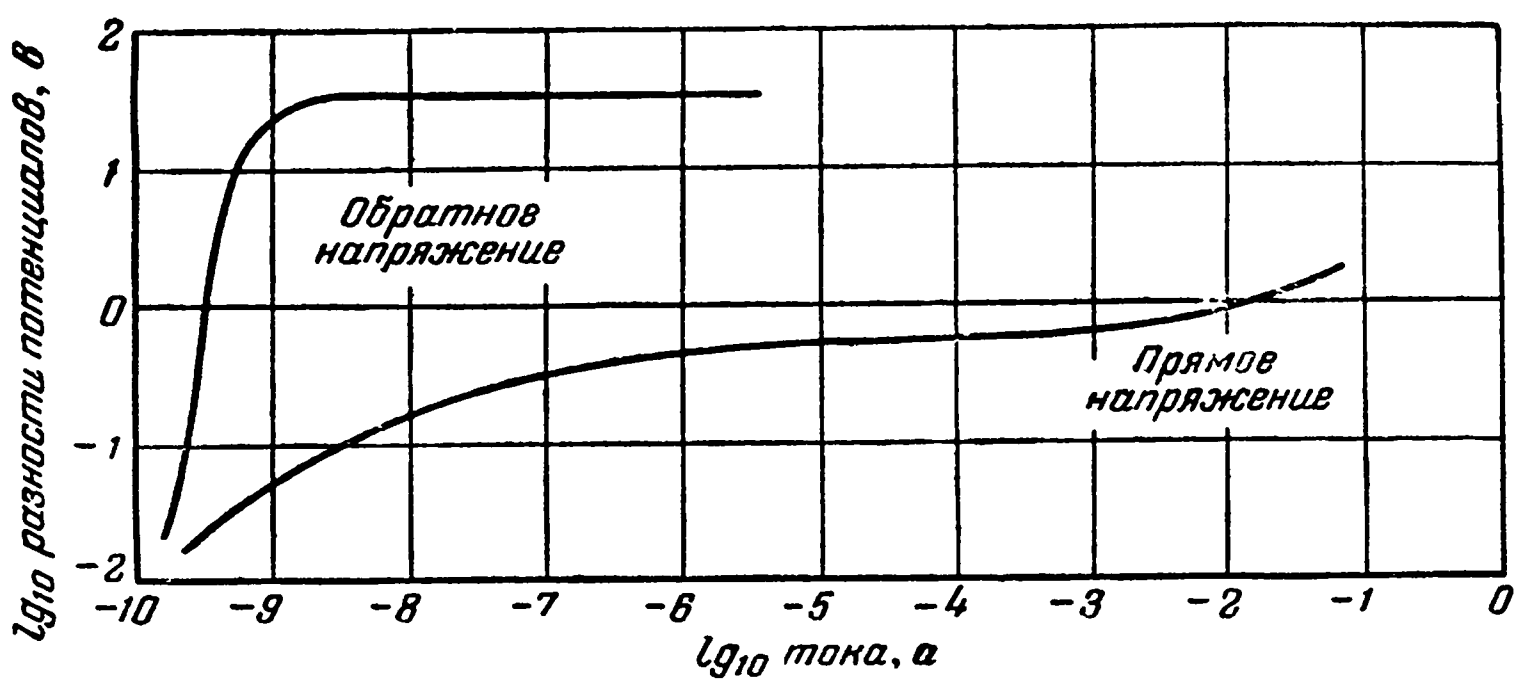


Рис. 6-16. Коэффициент выпрямления кремниево-алюминиевого сплавного диода может досигать 100 000 000 при напряжении 1 *в*.

Не следует думать, что такой высокий коэффициент выпрямления является результатом использования сплавного перехода вместо выращенного, хотя геометрические особенности такого диода, приближающегося к точечному, несколько улучшают характеристику в прямом направлении за счет эффекта распределенного сопротивления. Значительно бóльшую роль в улучшении коэффициента выпрямления сплавного диода играет использование кремния в качестве полупроводникового материала. Хорошо обоснованные физические причины объясняют, почему *p-n* переход в кремнии должен иметь обратное сопротивление, в 1 000—10 000 раз превышающее таковое для *p-n* перехода в германии при одинаковой площади контакта. Эти причины разобраны в гл. 19.

6-5-2. Эффекты, связанные с накоплением неосновных носителей

Физическим явлением, играющим роль в общей картине поведения выпрямителей при быстрых переходных процессах, оказывается задержка (запаздывание) на-

копления) неосновных носителей [Л. 6 и 7]. Хотя эти явления запаздывания не очень опасны при использовании кремниевых и германиевых точечных диодов при детектировании сигналов в диапазоне сантиметровых волн, однако они часто являются ограничивающим фактором при использовании диодов с *p-n* переходом в коммутационных цепях схем СВЧ.

Если через выпрямитель в течение нескольких микросекунд (или дольше) проходит постоянный прямой ток, после чего поданное напряжение внезапно меняет полярность с прямой на обратную, то требуется некоторое время, для того чтобы обратное сопротивление диода приняло величину, соответствующую статической характеристике диода. В течение этого времени установления диод пропускает сравнительно высокий обратный ток. На рис. 6-17 показаны графики напряжения и соответствующего тока в течение периода коммутации. Непосредственно после переключения полярности входного напряжения диод может пропускать обратный ток, сравнимый по величине с прямым током, протекавшим через диод в момент, непосредственно предшествовавший коммутации. Однако этот переходный выброс быстро спадает, и спустя короткое время обратный ток устанавливается на значении, соответствующем его статической характеристике.

При рассмотрении эквивалентной схемы выпрямителя можно ожидать переходного процесса такого рода, время установления которого равно постоянной rC емкости перехода, разряжающейся через предельное сопротивление массы диода. Однако нетрудно рассчитать, что для практически наблюдающихся величин r и C длительность переходного процесса не должна превышать нескольких тысячных микросекунды. В то же время для некоторых *p-n* переходов измеренные длительности переходных процессов имеют величины порядка десятой микросекунды, т. е. намного бóльшие, чем вызываемое влиянием емкости перехода.

Запаздывание переходного процесса происходит за счет накопления заряда неосновных носителей в полупроводнике вблизи выпрямительного перехода. Если к *p-n* переходу приложено некоторое прямое напряжение, то бóльшая часть прямого тока создается частицами, которые были основными носителями в областях их

образования, но стали неосновными после того, как они проникли через переход. Таким образом, прямой ток является в основном суммой двух отдельных токов: тока электронов, образующихся в *n*-области перехода и протекающих через него в *p*-область, и тока положительных дырок, образующихся в *p*-области и протекающих в *n*-область. При этом процессе на небольших расстояниях по

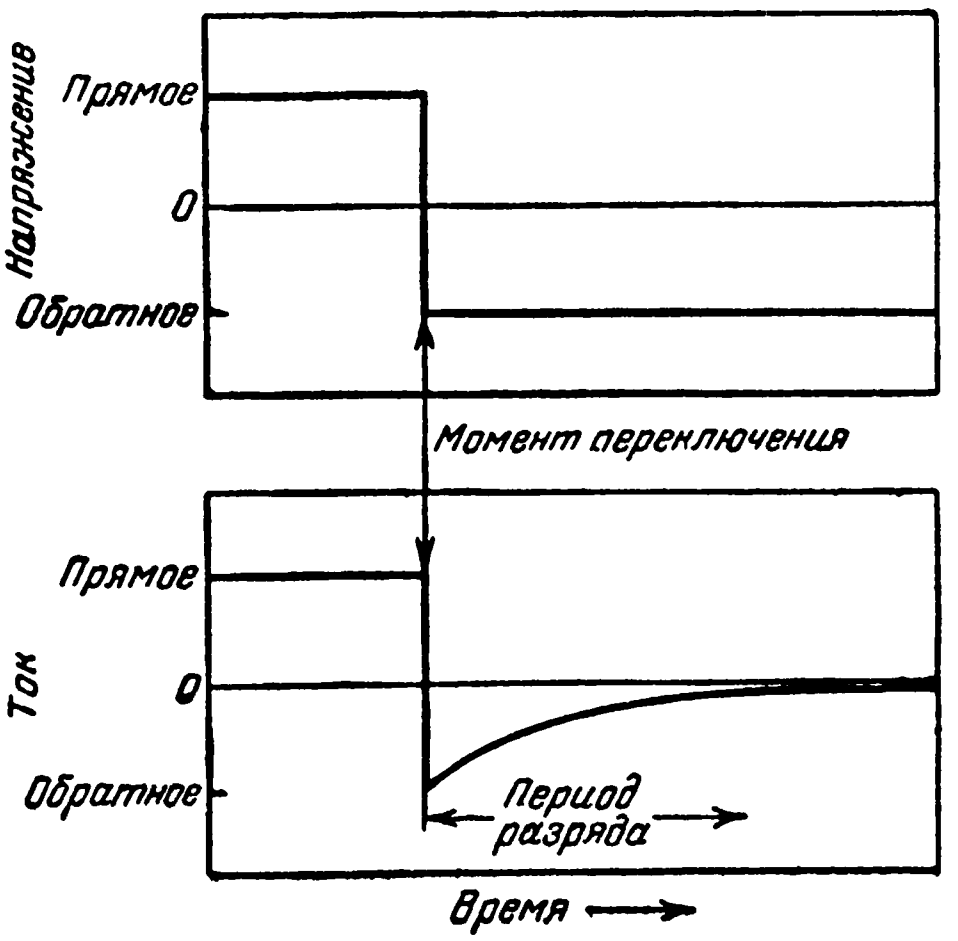


Рис. 6-17. При мгновенном переключении напряжения диода с прямого на обратное (вверху) в течение короткого интервала времени проходит большой обратный ток (внизу), прежде чем восстанавливается нормальное обратное сопротивление.

обе стороны от перехода поддерживаются избыточные концентрации неосновных носителей.

Если постоянный прямой ток шел достаточно долго, чтобы привести эти концентрации в состояние динамического равновесия, и если полярность входного напряжения внезапно меняется, то значительное количество этих «накопленных» неосновных носителей диффундирует через переход в обратном направлении, создавая относительно большие обратные токи. Эти токи проходят до тех пор, пока запасы накопленных неосновных носителей не исчерпываются частично за счет указанных токов, частично за счет рекомбинации на месте. Только тогда сопротивление переходного слоя принимает высокое значе-

ние, соответствующее действующему обратному напряжению.

Длительность времени задержки за счет накопления неосновных носителей сложным образом зависит от прямого смещения в момент, непосредственно предшествующий переключению, от времени жизни неосновных носителей в *n*- и *p*-областях выпрямителя, от подаваемых прямого и обратного напряжения и от общего сопротивления цепи выпрямителя. Эффект накопления не ограничивает работу диодов на высоких частотах в режиме детектирования сигналов, поскольку периоды подачи прямого напряжения при этом слишком коротки для образования значительных концентраций накопленных неосновных носителей перед быстро следующим очередным изменением полярности входного напряжения. Основные неприятности от эффекта накопления неосновных носителей сказываются на работе быстродействующих коммутирующих цепей, где бывает необходимо быстро разомкнуть диодный ключ после сравнительно длительного периода замкнутого состояния. Очевидным способом уменьшения задержки за счет накопления зарядов является выбор для выпрямителя полупроводника с малым временем жизни неосновных носителей. К сожалению, это связано с нежелательными изменениями других параметров; в частности, может понизиться обратное сопротивление выпрямителя. Зависимость между временем жизни неосновных носителей и обратным током насыщения выведена в гл. 19.

ЗАДАЧИ

1. Постройте график для полного сопротивления $Z=R_s+jX_s$ на комплексной плоскости для различных частот в пределах от 0 до ∞ для выпрямителя, имеющего постоянные величины R_{ac}, r и C. Рассмотрите форму полученной кривой и объясните ее особенности.

2. Точечный выпрямитель имеет круглый контакт диаметром 0,02 *мм* на полупроводнике с удельным сопротивлением 1 *ом·см*. Каково предельное прямое сопротивление этого выпрямителя? Приняв эффективную емкость перехода равной 0,05 *мкф/см*^{2}, определите предельную частоту этого выпрямителя.

3. Толщина выпрямительного слоя, рассчитанная из уравнения Шоттки (6-2) и величин емкостей, приведенных на рис. 6-3 и 6-4, очевидно, повышается пропорционально *квадратному корню* из обратного напряжения. Однако обратное сопротивление выпрямителя с переходом, как видно из характеристики на рис. 5-15, растет *линейно* с приложенным обратным напряжением в пределах напряжений, соответствующих току насыщения. Объясните это видимое несоответствие.

ЛИТЕРАТУРА

1. Schottky W., Vereinfachte und erweiterte Theorie der Randschichtgleichrichter, Zeitschrift für Physik, 1941—1942, Bd. 118, S. 539.
2. Torrey H. C., Whitmer C. A, Crystal Rectifiers, McGraw-Hill, New York, 1946.
3. Jeans J. H, Mathematical Theory of Electricity and Magnetism, Cambridge, 5th ed., 1925, p. 356
4. Pearson G. L, Foy P. W., Silicon p—n Junction Diodes Prepared by the Alloying Process, Physical Review, 1952, vol. 87, p. 190.
5. Pearson G. L., Sawyer B., Silicon p—n Junction Alloy Diodes, Proceedings of the Institute of Radio Engineers, 1952, vol. 40, p 1348.
6. Meacham L. A, Michaels S. E., Observations of the Rapid Withdrawal of Stored Holes from Germanium Transistors and Varistors, Physical Rewiew, 1950, vol. 78., p. 175.
7 Kingston R. H., Switching Time in Junction Diodes and Junction Transistors, Proceedings of the Institute of Radio Engineers, 1954, vol. 42, p. 829.

Веденеев Г. М. и Вершин В. Е., Кремниевые стабилитроны, ГЭИ, 1961.

Туннельные диоды, сборник статей под ред. В. И. Фистуля, Изд. иностранной литературы, 1961.

Кремниевые стабилитроны Д-808-Д 813, «Радио», 1959, 5, 61.

ГЛАВА СЕДЬМАЯ

ПРИМЕНЕНИЕ ВАРИСТОРОВ

Для полного описания областей применения варисторов, настолько подробного, чтобы мог понять среднеподготовленный читатель, нужна книга в сотни, если не тысячи страниц. В настоящей главе описаны лишь некоторые характерные случаи применения варисторов, иллюстрирующие широкие и многообразные возможности этих замечательных приборов. Как и ранее, все приводимые в тексте рисунки и графики служат только для пояснения принципов действия, но не руководством или справочным материалом для разработки практических схем.

7-1. УСТРОЙСТВА ДЛЯ ЗАРЯДА АККУМУЛЯТОРОВ

Одной из важных областей применения плоскостных выпрямителей является заряд аккумуляторов или поддержание аккумуляторов в заряженном состоянии. На рис. 7-1 показана простая цепь для зарядки аккумуляторов [1]. Такое устройство называется однополупериодным

[1] Стрелками показано прямое направление положительного тока.

выпрямителем, поскольку ток поступает в аккумулятор только в течение той части положительной полуволны входного напряжения, когда мгновенное напряжение источника питания превышает напряжение аккумулятора. Разряд аккумулятора через цепь заряда в течение отрицательной полуволны предотвращается высоким обратным сопротивлением выпрямителя, который в это время находится под обратным напряжением. Поэтому обратный разрядный ток оказывается пренебрежимо малым по сравнению с прямым током заряда. Таким образом,

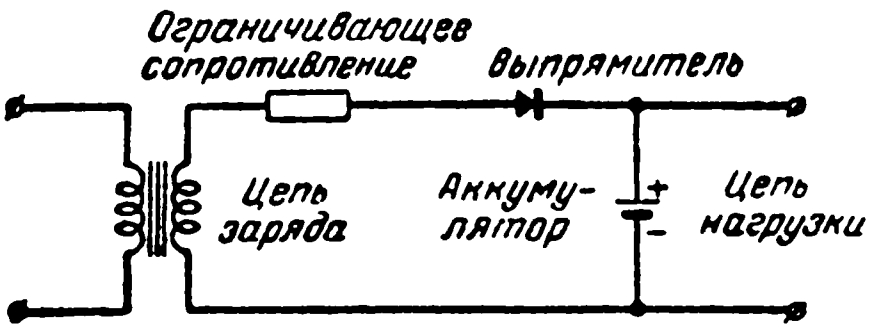


Рис 7-1 Полупроводниковые выпрямители широко применяются для зарядки аккумуляторов.

через аккумулятор проходит ток постоянный по направлению, но пульсирующий по величине.

Существенным при расчете зарядного устройства является требование, чтобы выпрямитель был способен выдержать обратное напряжение, равное сумме напряжения аккумулятора плюс амплитудное напряжение источника переменного тока. Если это условие не может быть удовлетворено с одним выпрямительным элементом, то нужно включить последовательно два или более выпрямительных элементов.

Большие зарядные установки снабжаются автоматическими регуляторами заряда, предназначенными для поддержания заряда батареи по мере ее разряда на рабочую цепь. Некоторые регуляторы выполняются в виде автоматов, которые включают батарею на заряд, когда она разряжена до определенной степени (что определяется по медленному падению напряжения по мере ее разряда), и снова отключают зарядное устройство, когда батарея полностью заряжена. Другие типы регуляторов обеспечивают при непрерывном заряде регулирование его интенсивности так, чтобы все время поддерживать батарею в заряженном состоянии.

7-2. СИЛОВЫЕ ВЫПРЯМИТЕЛИ [Л. 1, 2 и 3]

Для заряда аккумуляторных батарей вполне подходит однополупериодный тип выпрямителя, поскольку аккумулятор можно заряжать пульсирующим током постоянного направления или постоянным по величине током, имеющим то же среднее значение. Однако во многих случаях желателен более постоянный по величине ток, для чего можно создать различные схемы для

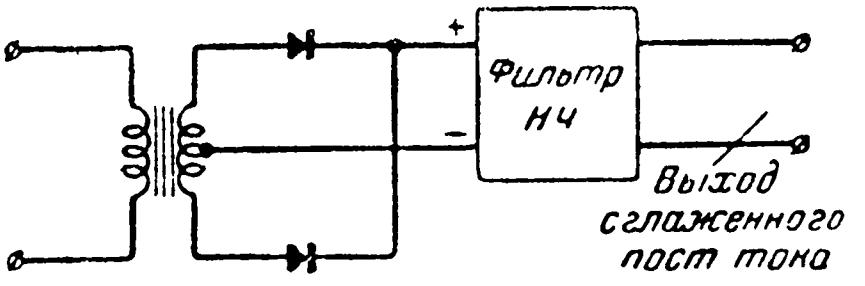


Рис 7-2 Двухполупериодный выпрямитель с двумя элементами и трансформатором с отводом от средней точки.

выпрямления обеих полуволн переменного тока. На рис. 7-2 показана одна из таких схем, в которой выпрямление производится двумя несимметричными вариторами, включенными в противоположных направлениях. В любой момент независимо от направления мгновенного напряжения питания один из выпрямителей заперт, а другой проводит ток, так что в выходной цепи ток проходит в течение обеих полуволн входного переменного напряжения. По этой причине такая схема называется двухполупериодным выпрямителем. Заметим, что выходное напряжение равно половине напряжения на вторичной обмотке трансформатора без падения напряжения на сопротивлении проводящего выпрямительного элемента. Однако максимальное обратное напряжение, которое должен выдерживать каждый выпрямительный элемент, равно полному напряжению вторичной обмотки без учета падения напряжения на другом элементе. Для выпрямления низких напряжений достаточно одного элемента в каждом плече такого выпрямителя. Если же требуется более высокое выходное напряжение, например для питания анодных цепей электронных ламп, вторичная обмотка должна давать соответственно более высокое напряжение и в каждом плече выпрямителя должно быть несколько последовательно

включенных элементов, способных вместе выдержать максимальное обратное напряжение.

Выходное напряжение двухполупериодного выпрямителя состоит из выпрямленных полуволн, в паузах между которыми ток и напряжение падают до нуля с частотой, равной удвоенной частоте входного напряжения. Если требуется постоянный по величине выходной ток, то выпрямленный ток пропускается через фильтр, который сглаживает пульсации и пропускает постоянный по величине усредненный ток. Тип этого фильтра определяет некоторые особенности конструкции выпрямителя;

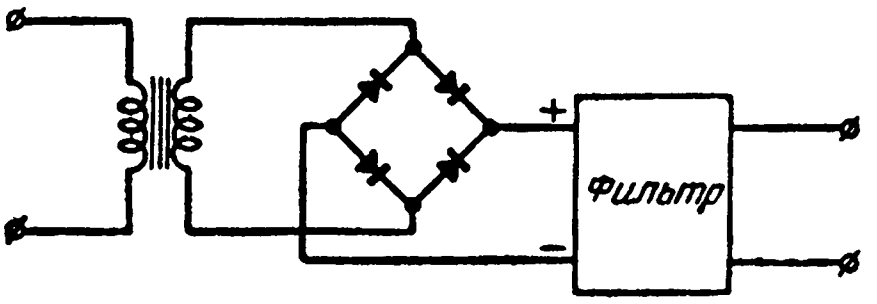


Рис. 7-3 Двухполупериодный мост из четырех выпрямительных элементов.

так, например, при фильтре с емкостью на входе обычно повышается максимальное обратное напряжение на элементах выпрямителя (вентилях) по сравнению с фильтром, имеющим на входе дроссель. Должны быть также приняты во внимание потери в фильтре при расчете напряжения вторичной обмотки трансформатора.

Варисторы часто включаются в схему четырехзвенного моста для двухполупериодного выпрямления (рис. 7-3). Мостовая схема выпрямителя имеет то преимущество, что она не нуждается в выводе средней точки от вторичной обмотки трансформатора. Недостаток этой схемы заключается в том, что если какая-либо точка входной цепи заземлена, то ни одну точку выходной цепи заземлять нельзя.

При расчете всех цепей выпрямителей исходят из заданной величины выходного напряжения, позволяющей определить входное напряжение и число выпрямительных элементов в каждом плече, необходимое для того, чтобы выдержать максимальное обратное напряжение. Желательный выходной ток определяет площадь выпрямительных элементов. В настоящее время расчет выпрямителей разработан подробно, практические правила изложены в справочниках по электротехнике и научно-технической литературе.

7-3. МОДУЛЯТОРЫ И ДЕМОДУЛЯТОРЫ

В системах уплотнения линий дальней связи необходимой операцией является модуляция [Л. 4 и 5] Под модуляцией подразумевается управление мгновенной величиной амплитуды, фазы или частоты непрерывных колебаний (несущей частоты) в соответствии с мгновенной величиной амплитуды управляющего напряжения, называемого сигналом. В системах передачи на несущей последняя генерируется в пункте передачи с помощью местного генератора, в то время как модулирующим сигналом является речь или другая информация, подлежащая передаче. Сигнал и несущая вводятся в модулятор, где и происходит модуляция; затем промодулированная несущая поступает на входной конец линии связи открытой проводной линии, коаксиального кабеля или радиорелейной линии Любое из этих технических средств передачи может одновременно передавать несколько различных несущих частот, каждая из которых промодулирована своим собственным сигналом. Таким образом, путем эффективного использования частотного спектра с помощью одного средства передачи можно передавать несколько сигналов одновременно.

Использование варисторов в качестве модуляторов возможно благодаря нелинейности их характеристик. В телефонии применение варисторов для указанных целей вместо электронных ламп имеет бесспорное экономическое преимущество. Не требуется источника энергии для накала ламп. Для оборудования, которое должно быть включено круглосуточно, но используется лишь в течение части этого времени, экономия за счет накала получается заметной. Кроме того, долговечность варисторов больше. В качестве модуляторов в соответствующих схемах могут работать как симметричные, так и несимметричные варисторы, но широкое применение нашел лишь последний вид варисторов; поэтому в дальнейшем мы рассматриваем только несимметричные приборы.

Чтобы понять, как происходит модуляция и как она применяется в многоканальной связи с частотным разделением, рассмотрим показанный на рис 7-4 случай идеально простой схемы модулятора, состоящей из выпрямителя, источника напряжения смещения U_0 и двух генераторов сигналов малой амплитуды. Напряжение

смещения ставит выпрямитель в режим, соответствующий некоторой статической рабочей точке, например P, на участке прямого тока его характеристики, где нелинейность имеет наиболее благоприятный характер для

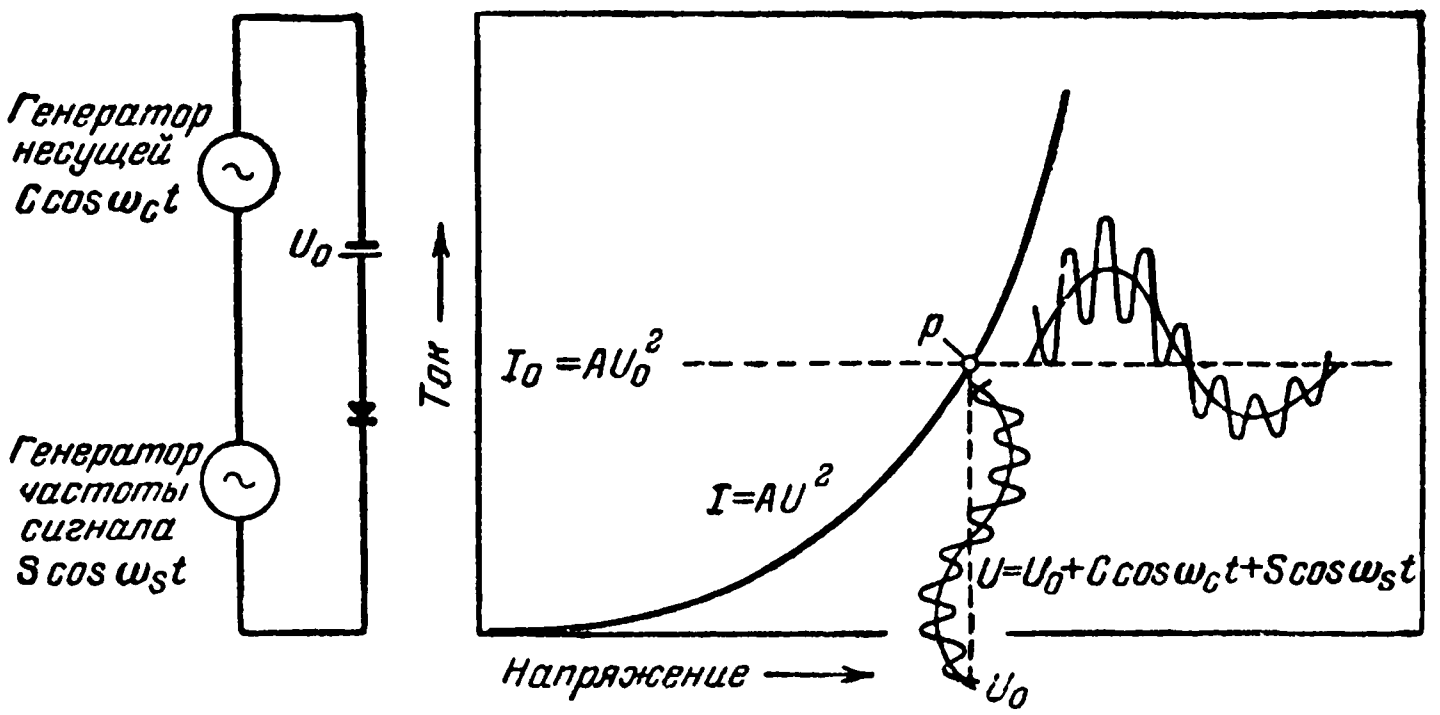


Рис. 7-4. В идеализированной цепи варисторного модулятора (слева) взаимодействуют синусоидальные напряжения, вырабатываемые генераторами колебаний несущей частоты и модулирующего сигнала, в результате чего получается кривая тока (справа), содержащая новые частотные составляющие.

описываемой цели. Для упрощения пренебрежем внутренними сопротивлениями источника напряжения смещения и генератора переменного тока и предположим, что вблизи точки P характеристика выпрямителя в точности квадратична:

$$I = AU^2. \tag{7-1}$$

Пусть один из генераторов (генератор несущей) генерирует переменное напряжение несущей частоты $f_c = \omega_c/2\pi$, тогда как другой, которым пусть будет микрофон, возбуждаемый голосом, создает гармонический сигнал звуковой частоты $f_s = \omega_s/2\pi$. На практике $\omega_c \gg \omega_s$. Мгновенное напряжение на выпрямителе тогда равно:

$$U = U_0 + C\cos\omega_c t + S\cos\omega_s t, \tag{7-2}$$

где C и S — амплитуды напряжений несущей и сигнала.

Рассмотрим, как выглядит кривая тока в цепи модулятора, и исследуем содержание частотных слагающих в ней. Возведя в квадрат выражение для напряжения

(7-2), подставим его в выражение для тока (7-1) и, упростив тригонометрические выражения, получим:

$$\left.\begin{aligned} I = A\left(U_0^2 + \frac{C^2}{2} + \frac{S^2}{2}\right); \\ + 2AU_0\,(C\cos\omega_c t + S\cos\omega_s t); \\ + \frac{A}{2}\left(C^2\cos 2\omega_c t + S^2\cos 2\omega_s t\right); \\ + 2CS\,[\cos(\omega_c + \omega_s)\,t + \cos(\omega_c - \omega_s)\,t]. \end{aligned}\right\} \quad (7\text{-}3)$$

Таким образом, мы видим, что выходной ток содержит слагающие частот, отличные от тех, которые вырабатываются двумя генераторами переменного тока. В частности, в нем содержатся: 1) постоянная составляющая; 2) слагающие входных частот[1] ω_c и ω_s; 3) слагающие удвоенных входных частот и 4) слагающие суммарной и разностной частот $(\omega_c+\omega_s)$ и $(\omega_c-\omega_s)$.

Подробный анализ для более реального случая цепи, содержащей последовательно включенное нагрузочное сопротивление и генераторы с конечными внутренними сопротивлениями, показывает, что полученные ранее результаты изменяются только в отношении амплитудных коэффициентов при различных членах выражения для тока, но не в отношении состава частотных слагающих, содержащихся в нем. Предположим, как показано на рис. 7-5, что в цепь введены выходной трансформатор и частотный фильтр, пропускающий только «верхнюю боковую» частотную слагающую ACS $\cos(\omega_c+\omega_s)t$. Тогда окажется, что мы перенесли первоначальную частоту сигнала ω_s в другой участок частотного спектра $\omega_c+\omega_s$, на частоте которого и происходит теперь передача информации.

Рассмотрим, как такой перенос частот позволяет производить одновременную передачу нескольких разговоров по одной паре проводов. Пусть имеется несколько сигналов низкой частоты с частотами f_{s1}, f_{s2}, f_{s3}.., создаваемые несколькими собеседниками, каждый из которых говорит на одной единственной частоте в свой микрофон. (На самом деле телефонные разговоры не ведутся только на одной частоте, но мы допускаем это

---

[1] Обычно под словом «частота» подразумевается как число периодов в секунду (частота f), так и соответствующая угловая частота ω В тексте этот термин используется в обоих значениях Из символов f или ω всегда ясно, о какой именно частоте идет речь

для простоты.) Пусть цепь каждого абонента подобна изображенной на рис. 7-5 и связана с общей линией при помощи трансформатора. На рис. 7-6 показана такая схема для трех абонентов. Схема в целом является передающей частью трехканальной системы уплотнения линии связи. Каждый генератор несущей создает частоту, отличную от двух остальных; каждый частотный фильтр настроен так, чтобы пропускать только верхнюю боко-

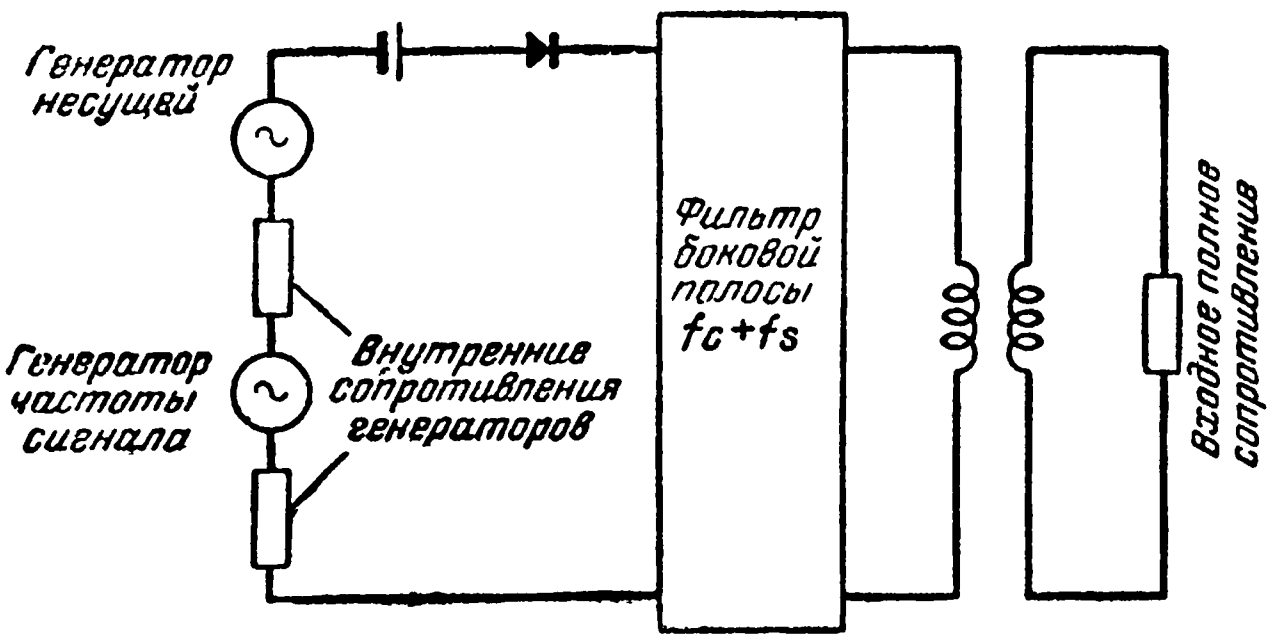


Рис. 7-5. В практически применяемом модуляторе приходится учитывать сопротивления различных элементов цепь, фильтр боковой полосы и величину связи с линией

вую полосу частот (f_c+f_s), соответствующую данной несущей и частоте сигнала. В линию поступают одновременно три информации, передаваемые на трех различных участках спектра частот ($f_{c1}+f_{s1}$; $f_{c2}+f_{s2}$; $f_{c3}+f_{s3}$).

На приемном конце нужно разделить эти три боковые полосы, перенести несущие информацию частоты снова в область звуковых частот и довести полученные низкочастотные сигналы до абонентов. Эти функции выполняются с помощью аппаратуры, подобной уже описанной.

Все основные особенности рассмотренной выше картины полностью сохраняются и в том случае, если варисторы, используемые в качестве модуляторов, вместо простой квадратичной характеристики имеют вблизи рабочей точки любую нелинейную характеристику, выражаемую функцией $I=AU^n$, где n — любое число, отличное от единицы. При более высоких целых значениях n выражение для тока, соответствующее уравнению (7-3), содержит высшие гармоники частот f_e и f_s,

а также суммарные и разностные частоты (hf_c+kf_s), где h и k — целые числа. Если n не целое число [1], как это обычно имеет место в практике для модулятора на варисторах, то в выражении для тока появится бесконечное число членов, из которых лишь немногие имеют достаточные амплитуды, чтобы представлять технический интерес. Описанная картина также существенно

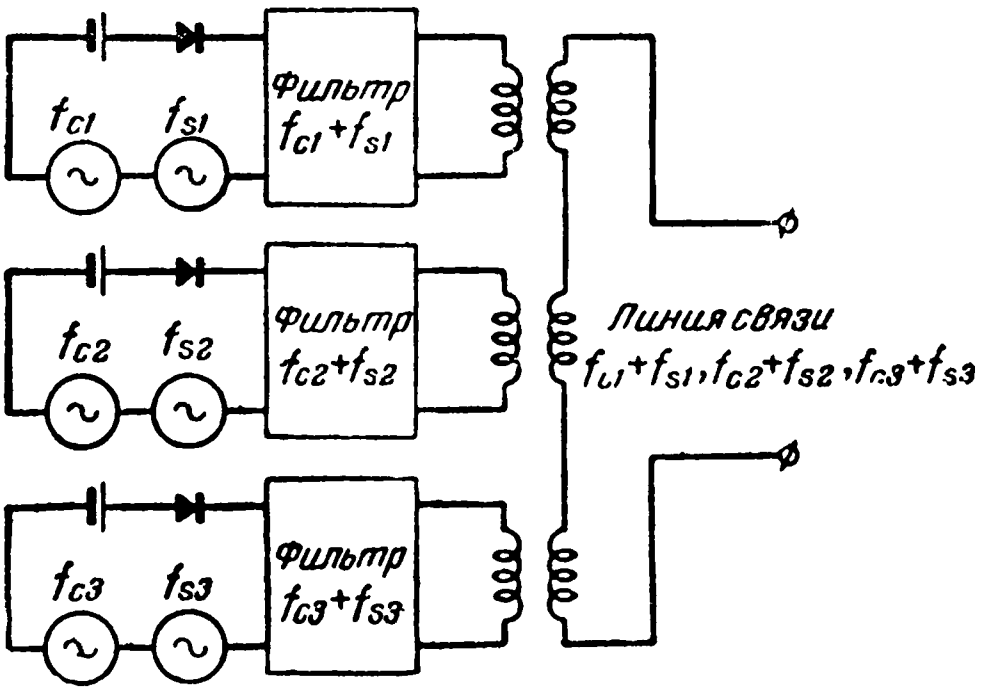


Рис. 7-6 Схема, позволяющая осуществить одновременную передачу разговоров трех абонентов: f_{s1}, f_{s2} и f_{s3} через одну и ту же линию связи без взаимных помех.

не меняется, если мы допустим, что каждый абонент пользуется вместо одной частоты набором звуковых частот от 100 до примерно 3 500 *гц*. Поэтому передаваемая боковая полоса частот, содержащая речь, например, абонента № 1, включает область частот от $f_{c1}+$ $+100$ *гц* до $f_{c1}+3\,500$ *гц*. Такая полоса частот называется «каналом» [2]. Одной из многих задач при разработке систем многоканальной проводной связи являются выбор несущих частот f_1, f_2, f_3... достаточно далеких друг от друга, создание избирательных фильтров и других схемных элементов, достаточно линейных, чтобы исключить взаимное перекрытие несущих информацию боковых частот. В противном случае возникают перекрестные помехи.

---

[1] Если n не целое число, то характеристику можно представить степенным рядом U. Такой анализ приведен в [Л. 6].

[2] Термином «канал» иногда обозначают также совокупность технических средств, используемых для передачи одной такой полосы звуковых частот

Можно сказать, что чем сильнее выражена нелинейность характеристики модулятора, тем больше амплитуды слагающих боковых полос в выходном сигнале модулятора. Поэтому для модуляторов предпочтительно применять варисторы с большой величиной показателя n.

В разобранном выше для иллюстрации примере рассмотрен случай модуляции при малых сигналах; однако большинство модуляторов в многоканальных телефонных системах работает в режиме больших сигналов

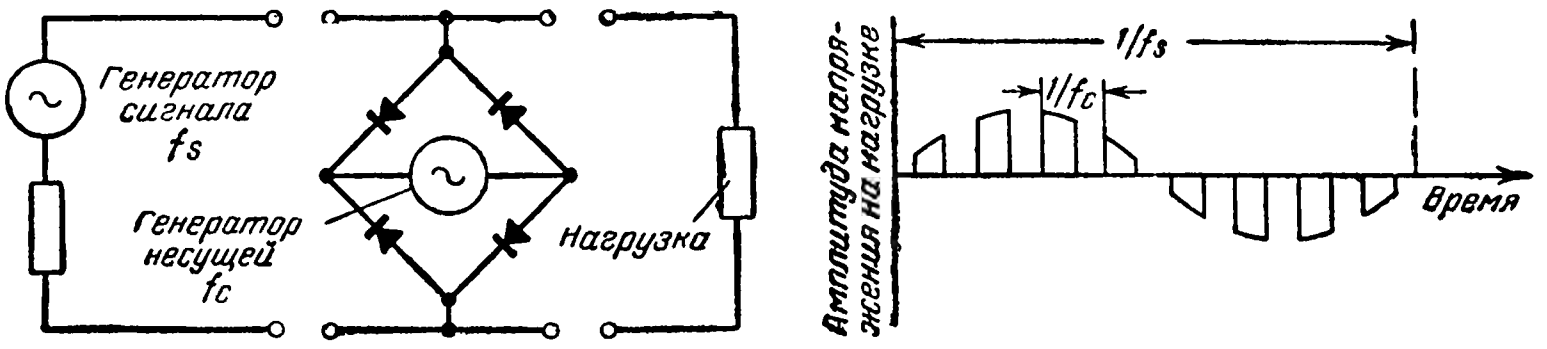


Рис 7-7. Мостовой модулятор на варисторах и форма сигнала, вырабатываемого им (справа).

[Л. 7]. На рис. 7-7 показана мостовая схема для модуляции большим сигналом с помощью четырех варисторов. Напряжение несущей частоты f_c от генератора вводится в одну диагональ моста, а модулирующий сигнал звуковой частоты f_s вводится в другую диагональ параллельно с линией передачи. В этом случае амплитуда сигнала несущей частоты имеет величину 1—2 *в*, тогда как модулирующий сигнал имеет амплитуду порядка нескольких милливольт. В течение половины периода несущей частоты через все четыре варистора проходит прямой ток. При этом мост имеет малое сопротивление для сигнала и бóльшая часть последнего проходит через мост, шунтирующий линию. В течение следующего полупериода несущей частоты варисторы запираются приложенным обратным напряжением и сигнал поступает в линию. Результирующая форма кривой выходного напряжения показана в правой части рис. 7-7; амплитуда каждого второго полупериода несущей частоты понижается практически до нуля.

Анализ этой кривой показывает, что она содержит гармоники несущей частоты и модулирующего сигнала, а также ряды суммарных и разностных частот ($hf_c \pm$

$\pm kf_s$). Одна из этих боковых полос выделяется затем с помощью фильтра для дальнейшей передачи.

Другая популярная схема модулятора больших сигналов на варисторах — кольцевой модулятор — показана на рис. 7-8. В этом модуляторе напряжение несущей частоты в течение полупериода отпирает одну пару варисторов и запирает другую. В течение следующего полупериода отпирается вторая пара, а первая запирается. Таким образом, схема работает как двухпозиционный переключатель на два направления, срабатывающий

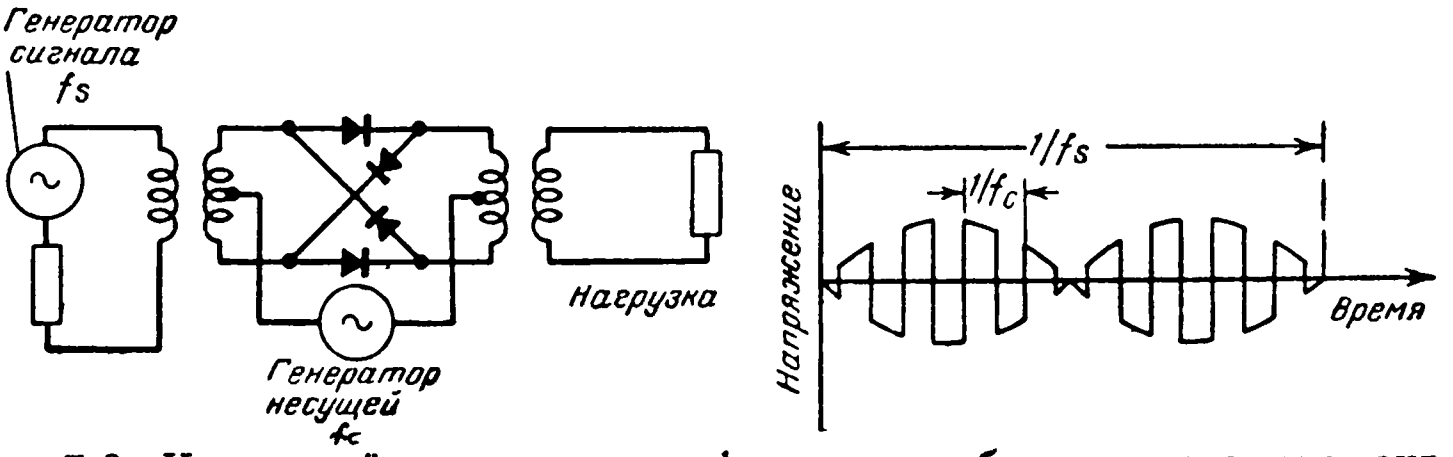


Рис. 7-8. Кольцевой модулятор и форма вырабатываемого им сигнала

с частотой несущей. Форма выходного напряжения показана на том же рис. 7-8.

Перед второй мировой войной широкое и почти исключительное применение в качестве модуляторов в различных телефонных системах имели меднозакисные варисторы. В послевоенные годы были разработаны кремниевые и германиевые точечные диоды, которые открыли путь для модуляции на более высоких частотах, чем с помощью меднозакисных элементов.

Модуляция и демодуляция ныне развились в сложные и специализированные отрасли науки. Полная разработка модулятора должна включать рассмотрение внутренних сопротивлений и нагрузок, уровней напряжений, модуляционных искажений, баланса, перекрестных помех и шумов. Однако эти вопросы относятся к области расчета цепей и систем, а не к области полупроводниковых приборов.

7-4. ЭКСПАНДЕРЫ И КОМПРЕССОРЫ

Передающая или воспроизводящая система может удовлетворительно работать лишь в ограниченной области амплитуд сигналов. Для высококачественной пере-

дачи уровень самых слабых сигналов должен быть много выше уровня шумов системы, а самые сильные сигналы должны все же оставаться ниже того уровня, где начинаются искажения, и не перегружать систему. Если попытаться точно передать в граммофонной записи весь диапазон громкостей сигналов, то либо громкость самых тихих пассажей окажется недопустимо близкой к уровню собственных шумов пластинки, либо для передачи громких пассажей потребуется экономически невыгод-

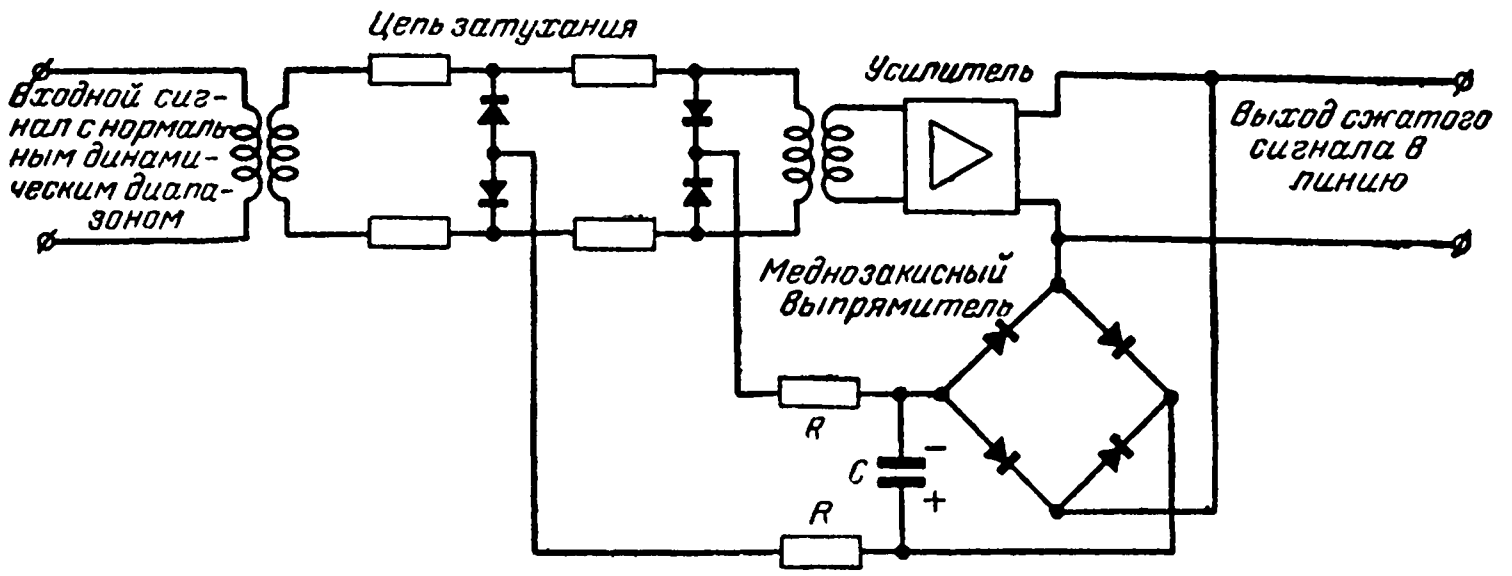


Рис. 7-9. Компрессор, осуществляющий сжатие динамического диапазона амплитуд сложного сигнала.

ное увеличение расстояния между канавками (шаг записи).

Пусть некоторая система имеет с учетом достаточного превышения сигнала над уровнем шумов динамический диапазон всего лишь 15 или 20 *дб*, в пределах которого ее амплитудные характеристики остаются линейными; все же может оказаться желательным с помощью этой системы записать или передать сигналы с динамическим диапазоном 40 или 50 *дб*. В этом случае система должна иметь устройства для сжатия (компрессии) динамического диапазона сигнала на входе и расширения (экспандирования) на выходе системы. Устройство для выполнения этих задач в зависимости от точного смысла выполняемых ими функций получили разные названия: компрессоры и экспандеры, ограничители громкости, усилители с корень-квадратной и квадратичной характеристиками. Для нашего обзора достаточно привести описание одного из таких устройств — «компандера» [Л. 8 и 9].

На рис. 7-9 показана схема с варисторами для сжатия амплитуд сигнала на входном конце линии, оборудованной «компандером». В этой схеме некоторая часть усиленного сигнала поступает на мостовой выпрямитель, который вырабатывает прямое напряжение для четырех варисторов в схеме затухания, управляя главным образом величиной затухания, вносимого этой схемой. В результате, осуществляется общая компрессия, так как сигналы с большой амплитудой вызывают большое за-

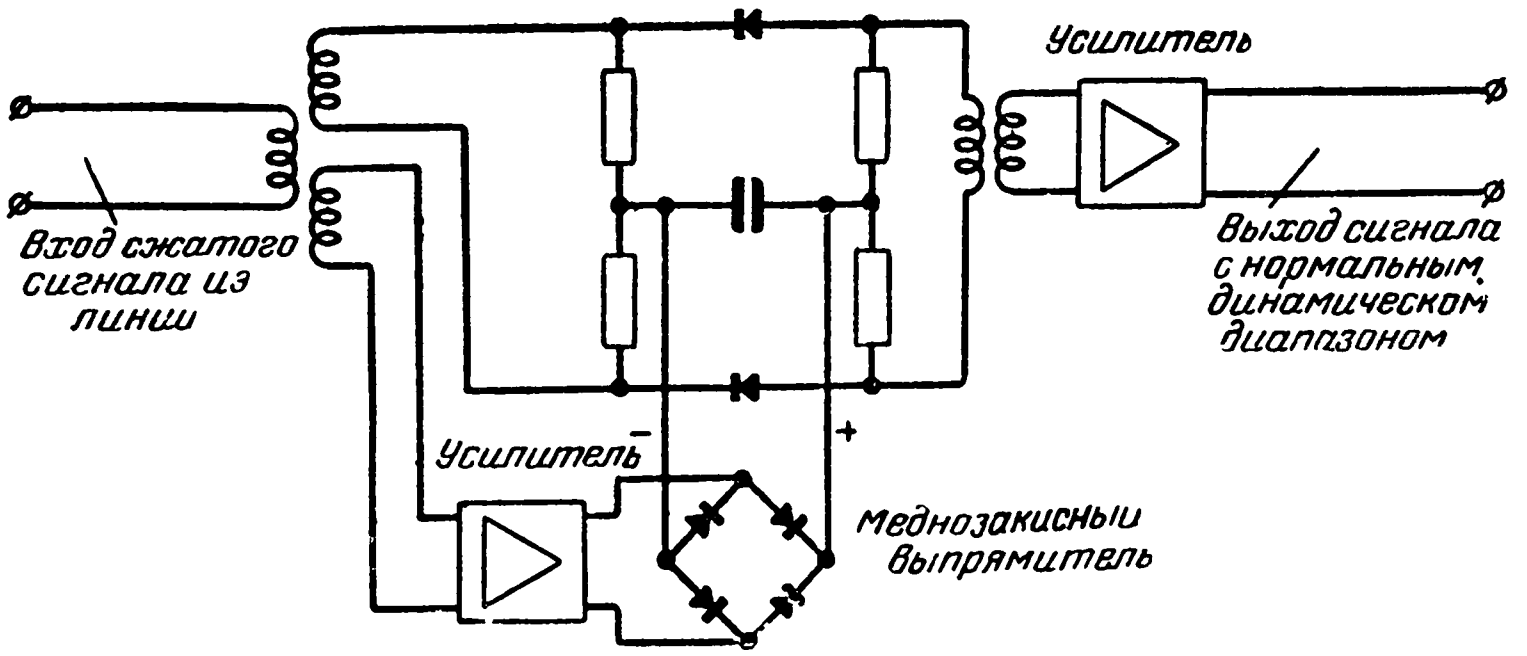


Рис. 7-10. Экспандер (расширитель), осуществляющий восстановление динамического диапазона сигнала, сжатого в устройстве, показанном на рис. 7-9

тухание, чем слабые сигналы, и величина затухания управляется амплитудой самого сигнала. В цепь обратной связи введен *RC*-фильтр для сглаживания постоянного напряжения, поступающего в схему затухания. Без этого схема реагировала бы на каждый отдельный полупериод сигнала, а не на его огибающую, что привело бы к искажениям сигнала. Неизбежное следствие такого сглаживания — некоторое запаздывание срабатывания компрессора при внезапном приходе громкого сигнала. По этой причине описанная схема носит название «слогового» компрессора.

На рис. 7-10 показана схема экспандера на приемном конце линии. Заметим, что варисторы в схеме затухания включены здесь не параллельно, а последовательно, что приводит к понижению затухания при повышении амплитуды сигнала, обеспечивая этим расширение динамического диапазона сигнала.

7-5. УСТРОЙСТВО ЗАЩИТЫ ОТ ПЕРЕГРУЗОК, ПОДАВИТЕЛИ КОММУТАЦИОННЫХ ПОМЕХ, ОГРАНИЧИТЕЛИ НАПРЯЖЕНИЯ И СХЕМЫ МОЛНИЕЗАЩИТЫ

Обладая свойством мгновенно изменять внутреннее сопротивление от большой до весьма малой величины при повышении прямого напряжения от менее 0,25 *в*, до более 0,5 *в*, несимметричные варисторы довольно широко применяют в устройствах защиты от перегрузок, ограничителях напряжения и подавителях коммутационных помех.

На рис. 7-11 слева приведена схема цепи, состоящей из пары несимметричных варисторов, включенных в про-

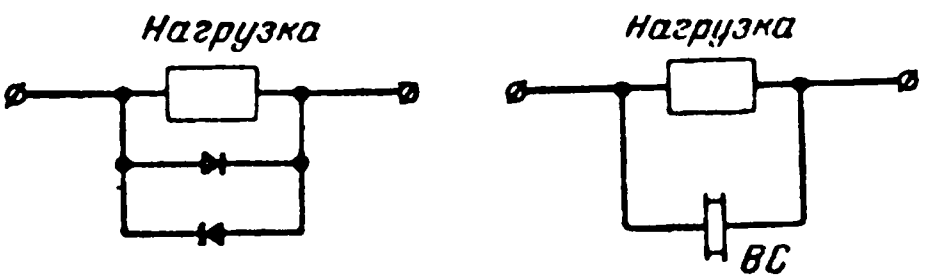


Рис. 7-11. Защита нагрузки от перенапряжений путем шунтирования парой несимметричных варисторов (слева) или одним симметричным варистором (справа).

тивоположных направлениях параллельно нагрузке, защищаемой от перегрузок. При сигналах малых амплитуд оба варистора имеют высокое сопротивление. Однако выброс напряжения любой полярности более чем на 0,5 *в* соответственно смещает рабочую точку варисторов и сопротивление того из них, для которого напряжение является прямым, сильно падает, в результате чего этот варистор в значительной степени шунтирует нагрузку.

Описанный способ широко используется для уменьшения нежелательных коммутационных помех, утомляющих слух телефонных абонентов, значительно ослабляя громкие щелчки, слышимые в телефонных трубках. Для этой цели используется пара меднозакисных выпрямителей диаметром 12 *мм*. Напряжение звуковой частоты с амплитудой порядка нескольких милливольт, подведенное к телефонной трубке сопротивлением 500 *ом*, лишь в незначительной степени шунтируется параллельной цепью защиты, которая при этом уровне напряжения имеет сопротивление около 1 500 *ом*. Однако при

возникновении коммутационной помехи с уровнем 1 *з* сопротивление одного из выпрямителей понижается примерно до 10 *ом*, обеспечивая, таким образом, отвод значительной части мощности помехи.

Симметричные варисторы подобным же образом применяются для защиты от перегрузок, как показано справа на рис. 7-11. Характеристики этих приборов позволяют использовать их для отделения полезных сигналов от помех при более высоких уровнях напряжения, чем это возможно с помощью пары несимметричных варисторов. Кроме того, использование симметричного варистора позволяет обойтись одним элементом вместо двух несимметричных элементов. Для защиты при напряжениях

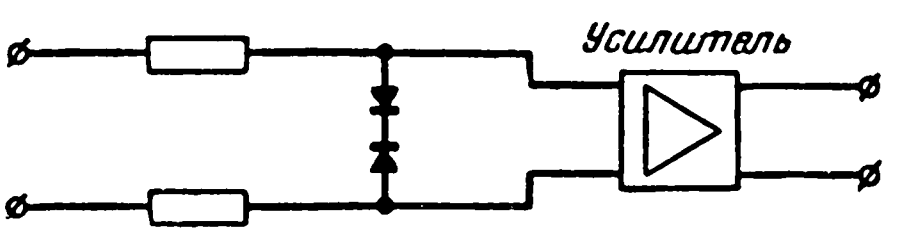


Рис 7-12 Защита усилителя от входных перенапряжений с помощью пары варисторов, работающих в режиме пробоя

в десятки и сотни киловольт, возникающих при грозовых разрядах, применяются симметричные тиритовые варисторы, включаемые между землей и линией параллельно защищаемой аппаратуре. Такие варисторы могут быть изготовлены в виде столбиков, составленных из ряда последовательно включенных тиритовых элементов диаметром ~ 10 *см*.

Относительной новинкой в области ограничения помех является выпрямитель с *p-n* переходом. Эффект ограничения возникает за счет крутого перегиба обратной характеристики на участке перехода от области насыщения с высоким сопротивлением к области обратного пробоя с малым сопротивление (рис. 5-15). Цепь ограничителя перегрузок, в котором используется явление пробоя, состоит из пары подобранных по величине сопротивления выпрямителей с *p-n* переходом, включенных последовательно навстречу друг другу, параллельно с прибором, подлежащим защите (рис. 7-12). Пока напряжение входного сигнала остается ниже напряжения пробоя любого выпрямителя, цепь защиты имеет сопротивление в десятки мегом. Однако для помех, амплитуда которых превышает обратное напряже-

ние пробоя в любом направлении, сопротивление этой цепи становится малым, шунтируя вход прибора и ограничивая излишнее напряжение помехи.

В этом случае диод, на который падает обратное напряжение, при перенапряжении попадает в режим пробоя, а второй диод оказывается под прямым напряжением и оба они вместе образуют шунт, замыкающий накоротко линию, прежде чем перенапряжение повредит защищаемое устройство. Мощность перенапряжения рассеивается при этом частью в диодах, частью в небольших включенных последовательно сопротивлениях и частью во входной линии.

Поскольку диоды с *p-n* переходом могут быть изготовлены с любым желаемым пробивным напряжением в широком практически используемом диапазоне, они весьма перспективны для защиты аппаратуры [Л. 11]. Устройства с пробивными диодами в настоящее время применяются для защиты промежуточных усилителей на транзисторах в системах многоканальной телефонной связи.

7-6. УСТРОЙСТВА ДЛЯ ЗАЩИТЫ КОНТАКТОВ РЕЛЕ

В момент прерывания тока в обмотке реле или какого-либо другого индуктивного элемента цепи на контактах выключателя (рис. 7-13 — слева) возникает значи-

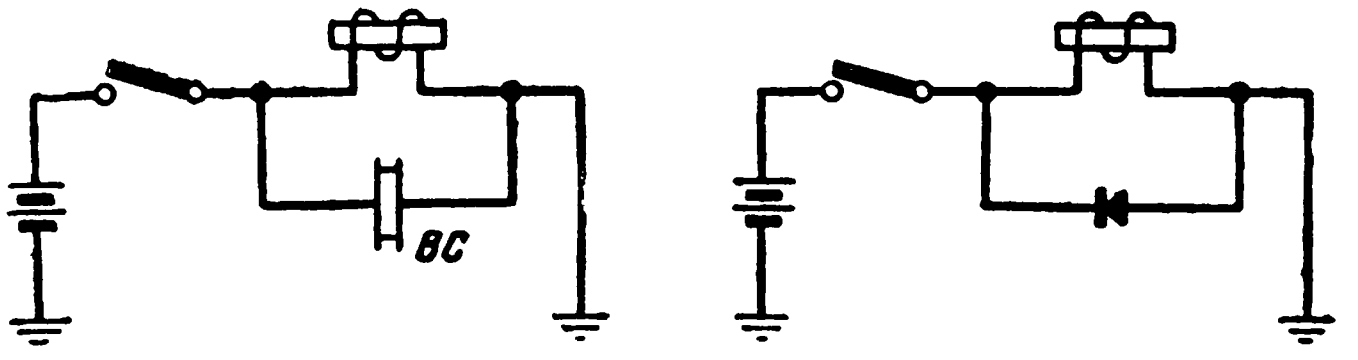


Рис. 7-13 Защита контактов от повреждения экстратоками в цепях, содержащих индуктивность, посредством шунтирования индуктивных элементов симметричными (слева) или несимметричными (справа) варисторами.

тельное напряжение. Этот бросок напряжения вызывает искрение, эрозию и разрушение контактов выключателя. Бóльшая часть таких вредных явлений устраняется путем включения надлежаще выбранного симметричного варистора параллельно обмотке. Когда через обмотку идет ток, то лишь пренебрежимо малая доля его ответ-

вляется в варистор. При размыкании выключателя индуцируемое в обмотке напряжение понижает сопротивление варистора до малой величины, создавая таким образом обходный путь для экстратока размыкания.

Для защиты контактов может быть использован и несимметричный варистор (рис. 7-13 — справа). Варистор включен в таком направлении, что напряжение на обмотке прибора является для него обратным; однако при размыкании тока полярность напряжения, наводимого в обмотке, оказывается для варистора «прямой», что позволяет рассеять мощность экстратока размыкания на сопротивлении варистора и обмотки без вреда для контактов. Этот же принцип применяется при несравненно бóльших токах и напряжениях и для защиты аппаратуры переменного тока при коммутационных переходных процессах. В силовых цепях, в состав которых входят обмотки генераторов, моторов или трансформаторов, при размыкании может возникнуть высокое индуцированное напряжение, сопровождаемое искрением и эрозией контактов выключателя или разрушением изоляции. В этих случаях для ограничения величины броска напряжения при размыкании применяются тиритовые варисторы подходящих размеров и характеристик [Л. 12 и 13].

7-7. ОГРАНИЧИТЕЛИ АМПЛИТУДЫ СИГНАЛОВ ПЕРЕМЕННОГО ТОКА

Радиоприемники сигналов частотной модуляции наиболее помехоустойчивы в том случае, когда остаточная амплитудная модуляция сигнала полностью подавляется. Хотя некоторые типы частотных детекторов теоретически должны быть чувствительны лишь к частотной модуляции, а не к амплитудной, их настройка для обеспечения нечувствительности к амплитудной модуляции оказывается весьма критичной. Поэтому часто в приемниках сигналов частотной модуляции предусматриваются ограничительные цепи, уменьшающие влияние остаточной амплитудной модуляции.

Если в каком-либо каскаде приемника, где сигнал имеет подходящую амплитуду, включить несимметричный варистор последовательно с источником питания так, чтобы он шунтировал цепь высокой частоты (рис. 7-14— слева), то пока амплитуда высокочастотного сигнала

меньше напряжения источника смещения, шунтирующее влияние цепи на сигнал пренебрежимо мало. Когда амплитуда сигнала несколько увеличится и превысит напряжение смещения, то диод начнет проводить и шунтировать цепь, срезая части кривой сигнала, превышающие напряжение смещения. Можно сделать симметричный ограничитель, если параллельно первой цепи включить вторую такую же цепь, состоящую из варистора и источника напряжения смещения, но включенную с противоположной полярностью. В этом

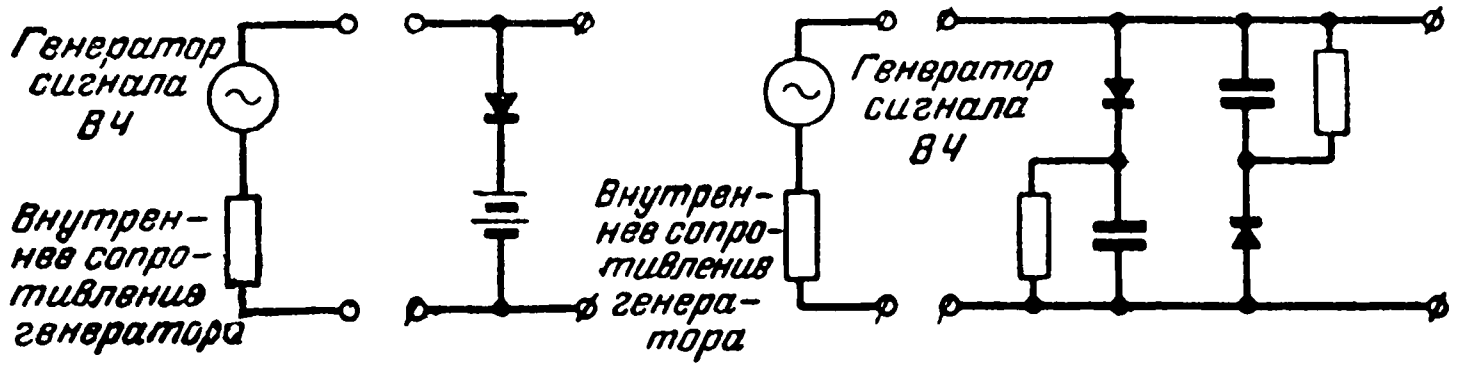


Рис 7-14 Ограничители амплитуды сигнала с фиксированным смещением (слева) или с автоматическим смещением, получаемым с помощью *RC* ячеек (справа).

случае ограничение происходит путем отсечки как верхних, так и нижних частей кривой сигнала, превышающих напряжение смещения.

Рассмотренные ограничители совершенно не действуют при напряжениях входных сигналов, меньших напряжения смещения. Однако при частотно-модулированном приеме существенным требованием является подавление амплитудной модуляции сигналов любой величины. Ограничитель, удовлетворяющий этому требованию, можно создать на базе описанных выше схем, заменив постоянные источники смещения *RC*-ячейками автоматического смещения, как показано на рис. 7-14 справа. Порог ограничения устанавливается автоматически в зависимости от амплитуд приходящего сигнала. Разумеется, постоянная времени ячейки *RC* должна быть достаточно велика по сравнению с периодом самой низкой частоты амплитудной модуляции, которую желательно подавить.

Имеется множество других схем использования варисторов как ограничителей амплитуды, подобных описанным выше. В системах импульсной модуляции варисторы используются для формирования прямоугольных

импульсов путем отсекания верхушек. Таким же образом варисторы используются в генераторах прямоугольных импульсов, формируемых из синусоиды отсечкой верхних и нижних частей полупериодов.

7-8. ДИОДНЫЕ ЛОГИЧЕСКИЕ СХЕМЫ

Молодой, но многообещающей областью применения варисторов, о которой придется очень много слышать в будущем, являются диодные логические схемы [Л. 14 и 15]. В качестве примера на рис. 7-15 приведен один из вариантов схемы типа И. Схема типа И имеет два (или

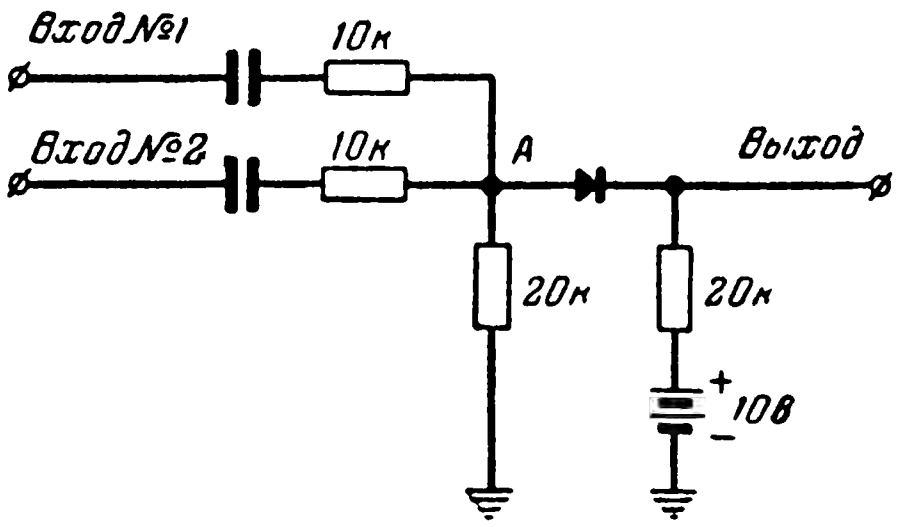


Рис. 7-15 Схема цепи типа «И» на диодах, в которой выходной импульс вырабатывается только в случае одновременного прихода двух импульсов на входы

больше) входов и один выход и построена так, что импульс на выходе появляется только в том случае, если импульсы на оба входа проходят одновременно. В такой схеме диод в обычном состоянии заперт обратным напряжением. Если на один из входов приходит импульс 20 *в*, причем второй вход остается заземленным, то напряжение в точке *А* схемы смещается в положительную сторону примерно на 8 *в*, причем диод остается запертым и на выходе ничего не появляется. Однако если на оба входа приходят импульсы 20 *в*, то напряжение в точке *А* смещается в положительную сторону примерно на 12 *в*, диод отпирается и на выходе появляется импульс с амплитудой примерно 2 *в*. Указанные напряжения, конечно, являются приблизительными; их точные величины зависят не только от параметров схемы и характеристик диода, но и от внутренних сопротивлений источников импульсов и выходной нагрузки.

Эта и подобные ей схемы, выполняющие такие функции, как «И», «НЕ», «И ИЛИ», «ЗАПОМИНАНИЕ», «ЗАПРЕЩЕНИЕ» и др., в которых используются диоды совместно с пассивными элементами цепи, образуют те «кирпичи», из которых состоит большинство элементов многих современных электронных вычислительных машин и других логических систем [Л. 16]. Рассмотрение математической основы этих операций, так же как и схем и приборов, с помощью которых они осуществляются, не входит в задачи нашей книги.

ЗАДАЧИ

1. На приведенном ниже рис. 7-16 пять селеновых выпрямительных элементов площадью 1 *см*^{2} каждый включены последовательно с нагрузкой 200 *ом* в схеме выпрямления переменного тока напряжением 50 *в*. Для упрощения расчетов предполагается, что источник

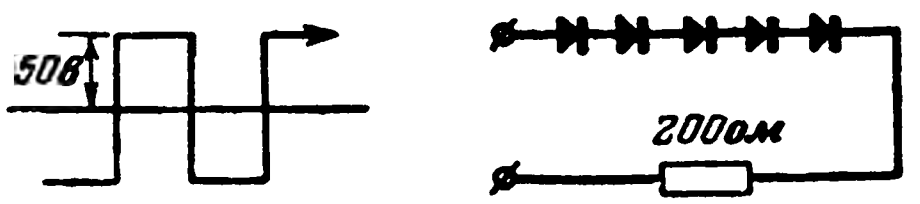


Рис. 7-16. К задаче 1.

дает не синусоидальное, а прямоугольное напряжение, как показано на графике. Определите для положительного полупериода: а) величину тока в цепи; б) напряжение на нагрузке; в) отношение потерь мощности в выпрямителе к мощности, выделяемой в нагрузке. (Ответы: а) 210 *ма*; б) 42 *в*; в) около 18%.)

2. В предыдущей задаче замените пять селеновых выпрямительных элементов одним кремниевым диодом площадью 1 *мм*^{2} с пере-

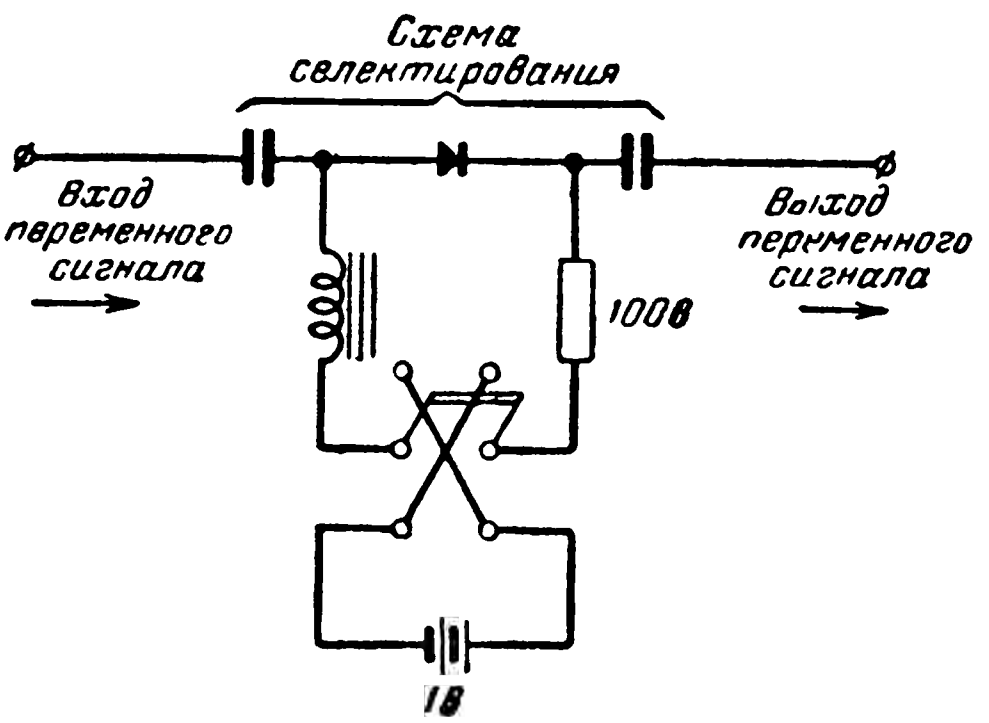


Рис. 7-17. К задаче 3.

менной проводимостью. Повторите расчеты и сделайте выводы из сравнения типов выпрямительных элементов

3. Меднозакисный выпрямитель, характеристики которого приведены на рис. 5-3, включен согласно приведенной на рис. 7-17 схеме селектирующего устройства для малых сигналов с управлением по постоянному току. Для обоих положений показанного на схеме переключателя найдите: а) ток и напряжение постоянного смещения на выпрямителе; б) сопротивление цепи переменному току. Допустите, что дроссель имеет бесконечное сопротивление переменному току, а конденсатор — бесконечную проводимость.

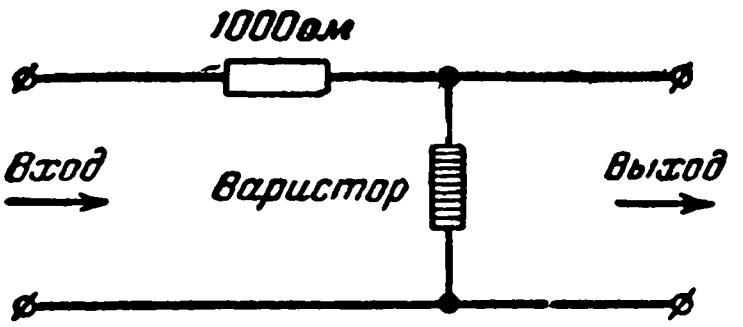


Рис. 7-18. К задаче 4.

4. Изготовлен дешевый и грубый стабилизатор постоянного напряжения с помощью варистора, включенного согласно приведенной на рис. 7-18 схеме. Варистор имеет самую нижнюю характеристику из трех, приведенных на рис. 5-21. Каково напряжение холостого хода стабилизатора для номинального входного напряжения 20 *в*? Во сколько раз изменится выходное напряжение холостого хода, если входное напряжение изменится вдвое (в любую сторону).

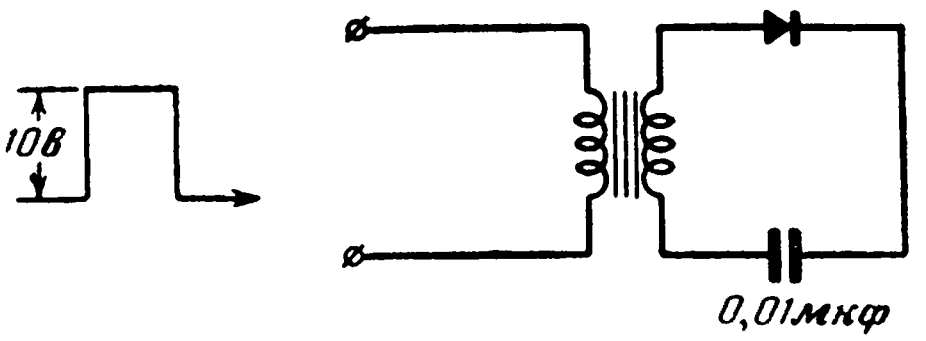


Рис. 7-19. К задаче 5.

5. Приведенная на рис. 7-19 цепь представляет собой схему блока «памяти» электронной вычислительной машины. Для запоминания «единицы» к первичной обмотке трансформатора с отношением обмоток 1 : 1 подводится импульс амплитудой 10 *в*. В течение этого импульса конденсатор заряжается до 10 *в* через прямое сопротивление кремние-алюминиевого диода (рис. 6-16). По окончании импульса заряд удерживается на конденсаторе за счет высокого обратного сопротивления диода. Определите приближенно, за сколько времени заряд конденсатора уменьшится на 10%, стекая через обратное сопротивление диода, если за это время заряд не будет стерт или возобновлен.

6. Симметричный варистор, имеющий самую нижнюю из изображенных на рис. 5-21 характеристик, включен для защиты от перегрузок параллельно входу усилителя, обладающего высоким

входным сопротивлением (рис. 7-20). Пусть в линии появляется помеха напряжением 500 *в*. Каково напряжение помехи на входе усилителя? Каково напряжение сигнала на входе усилителя, если нормальное напряжение переменного сигнала в линии 2 *в*?

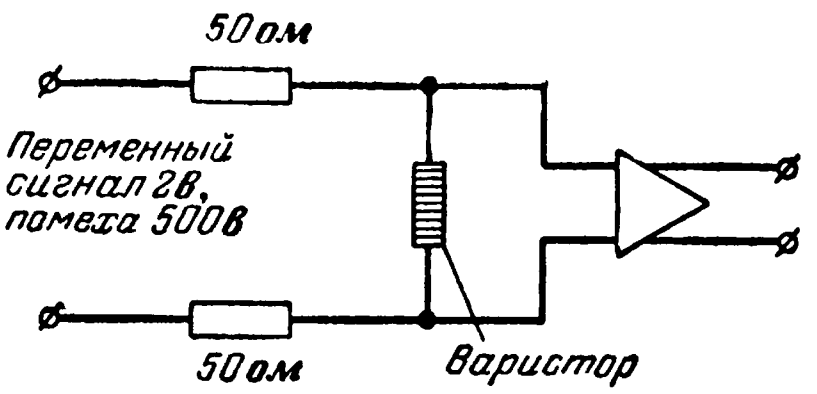


Рис. 7-20. К задаче 6

ЛИТЕРАТУРА

1. Clarke C. A., Selenium Rectifier Characteristics, Applications and Design Factors, Electrical Communication, 1941—1942, vol. 20, p. 47.

2. Beranek L. L., Applications of Copper Oxide Rectifiers, Electronics, 1939, vol. 12, p. 15.

3. Gramels J., Selenium Rectifiers—Factors in Their Application, Bell System Technical Journal, 1953, vol. 32, p. 1469.

4. Cowan F. A., Modulation in Communication, Transactions of the American Institute of Electrical Engineers, 1947, vol. 66, p. 792.

5. Terman F. E., Radio Engineers Handbook, McGraw-Hill, New York, 1943.

6 Peterson E, Llewellyn F. B, Operation of Modulators from a Physical Point of View, Proceedings of the Institute of Radio Engineers, 1930, vol. 18, p. 38.

7. Caruthers R. S., Copper Oxide Modulators in Carrier Telephone Systems, Bell System Technical Journal, 1939, vol. 18, p. 315.

8. Mathes R. C., Wright S. B., The Compandor, an Aid Agains Static in Radio Telephony, Electrical Engineering, 1934, vol. 53 p. 860; Bell System Technical Journal, vol. 13, p. 315.

9 Bennett W. R, Doba S., Vario-Losser Circuits, Electrical Engineering, 1941, vol. 60, p. 17.

10. Carter C. W., Dickieson A. C., Mitchell D., Application of Compandors to Telephone Circuits, Transactions of the American Institute of Electrical Engineers, 1946, vol. 65, p. 1079.

11. Pearson G. L., Silicon p—n Junction Power Rectifier and Lightning Protectors, Proceedings of the Institute of Radio Engineers, 1955, vol. 43, p. 487.

12. Albright J. W., Thyrite Protection for Series Windings of Autotransformers, Transactions of the American Institute of Electrical Engineers, 1955, vol. 74, part III, p. 150.

13. Meador J. R., Use of Thyrite in Power Transformers, Transactions of the American Institute of Electrical Engineers, 1949, vol 68, part II, p. 959.

14. Chen T. C., Diode Coincidence and Mixing Circuits in Digital Computers, Proceedings of the Institute of Radio Engineers, 1950, vol. 38, p. 511.

15. Hussey L. W., Semiconductor Diode Gates, Bell System Technical Journal, 1953, vol. 32, p. 1137.

16. Yokelson B. J., Ulrich W., Engineering Multi-Stage Diode Logic Circuits, Transactions of the American Society of Electrical Engineers, 1955, vol. 74, part I, p. 466

17. Полупроводниковые приборы с отрицательным сопротивлением, сборник статей под ред. С. А. Горяинова, Госэнергоиздат, 1962.

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА ПО ВАРИСТОРАМ

В перечисленных ниже работах, опубликованных за последние два десятилетия, приведены некоторые добавочные сведения об отдельных типах варисторов и их применении.

Angello S. J., Semiconductor Rectifiers, Electrical Engineering, 1949, vol. 68, p. 865.

Benzer S., High Inverse Voltage Germanium Rectifiers, Journal of Applied Physics, 1949, vol. 20, p. 804.

Richards E. A., The Characteristics and Applications of the Selenium Rectifier, Journal of the Institute of Electrical Engineers, 1941, vol. 88, part II, p. 423; vol. 88, part III, p. 238.

Scaff J. H., Ohl R. S., Development of Silicon Crystal Rectifiers for Microwave Radar Receivers, Bell System Technical Journal, 1947, vol. 26, p. 1.

Smith I. R, The Copper Oxide Rectifier, Electrical Engineering, 1948, vol. 67, p. 1051.

Stansel F. R., The Characteristics and Some Applications of Varistors, Proceedings of the Institute of Radio Engineers, 1951, vol. 39, p. 342.

Williams A. L., Thompson L. E, Metal Rectifiers, Journal of the Institute of Electrical Engineers, 1941, vol. 88, part I, p. 353.

ГЛАВА ВОСЬМАЯ

ПОЛУПРОВОДНИКОВЫЕ ФОТОЭЛЕМЕНТЫ

У полупроводников наблюдаются различные фотоэлектрические свойства (фотоэлектронная эмиссия, внутренний фотоэффект, фотогальванический эффект), рассмотренные в гл. 2; поэтому они в принципе могут служить для изготовления различных полупроводниковых приборов, действие которых основано на этих явлениях. Однако полупроводниковые фотоэмиссионные элементы практически еще недостаточно разработаны, чтобы конкурировать с электровакуумными фотоэлементами с металлическими катодами, используемыми в промыш-

ленных масштабах в течение примерно 30 лет [1]. С другой стороны, фотосопротивления и фотогальванические элементы принадлежат почти исключительно к полупроводниковому типу. Поэтому в настоящей главе рассмотрены именно эти приборы и описаны особенности их поведения.

8-1. ФОТОСОПРОТИВЛЕНИЯ

Основной частью фотосопротивления является полупроводниковый элемент, расположенный так, что на него может падать свет, и снабженный выводами, благодаря чему к нему может быть подведено напряжение и пропущен ток. На рис. 8-1 приведена схема такого простого полупроводникового фотосопротивления. Когда свет падает на полупроводник, его проводимость увеличивается и соответственно изменяется величина тока в присоединенной цепи. Механизм такого внутреннего фотоэффекта заключается, кратко говоря, в освобождении избыточных электронов и положительных дырок в толще полупроводника за счет поглощения световой энергии. Эти избыточные заряды повышают проводимость полупроводникового элемента пропорционально отношению их числа к числу избыточных частиц, имевшихся в полупроводнике до облучения его светом [2].

Освобождение новых электронов и дырок продолжается в течение всего времени освещения элемента. Однако быстрота рекомбинации электронов и дырок растет по мере повышения их концентрации, так что вскоре после начала освещения концентрация частиц достигает установившегося динамического равновесия, которое сохраняется, пока освещение не меняется.

После прекращения освещения избыточные частицы рекомбинируют друг с другом и восстанавливается прежняя величина проводимости, характерная для необлучаемого элемента. Время нарастания и спада концентра-

---

[1] Используемые в течение последних 20 лет электровакуумные фотоэлементы имеют не металлические катоды, а так называемые сложные фотокатоды, являющиеся по существу полупроводниковыми. *Прим. ред.*

[2] Этот вид внутреннего фотоэффекта называется положительным. Возможен отрицательный внутренний фотоэффект, при котором свет вызывает уменьшение электропроводности полупроводника, стимулируя захват и рекомбинацию возбужденных электронов. *Прим. ред.*

ции избыточных частиц зависит от времени жизни неосновных носителей заряда в данном полупроводниковом материале. В используемых на практике фотоэлементах время нарастания и спада изменяется от одной до нескольких сотен микросекунд и ограничивает наивысшую частоту изменения светового потока, при которой чувствительность фотоэлемента еще не начинает падать.

Если бы полупроводник имел вид толстого слоя, как показано на рис. 8-1, то увеличение проводимости за

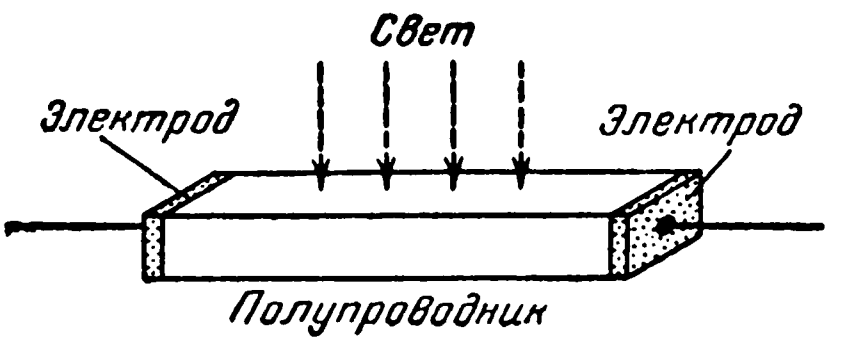


Рис. 8-1. Фотосопротивление состоит из полупроводника с электродами для подвода напряжения и устройством для освещения

счет внутреннего фотоэффекта имело бы место только в приповерхностных слоях, где происходит поглощение света. Хотя некоторые из освобождающихся электронов и дырок диффундируют в глубь полупроводника и воздействуют на проводимость слоев, лежащих глубже области проникновения света, однако глубина этой диффузии составляет лишь доли миллиметра, даже у тех полупроводников, у которых она наибольшая. Поэтому, для того чтобы получить наибольшее возможное относительное приращение проводимости фотоэлемента при данном световом потоке, слой полупроводника должен быть тонким. Выполнение этого требования составляет одну из трудностей конструирования фотосопротивлений.

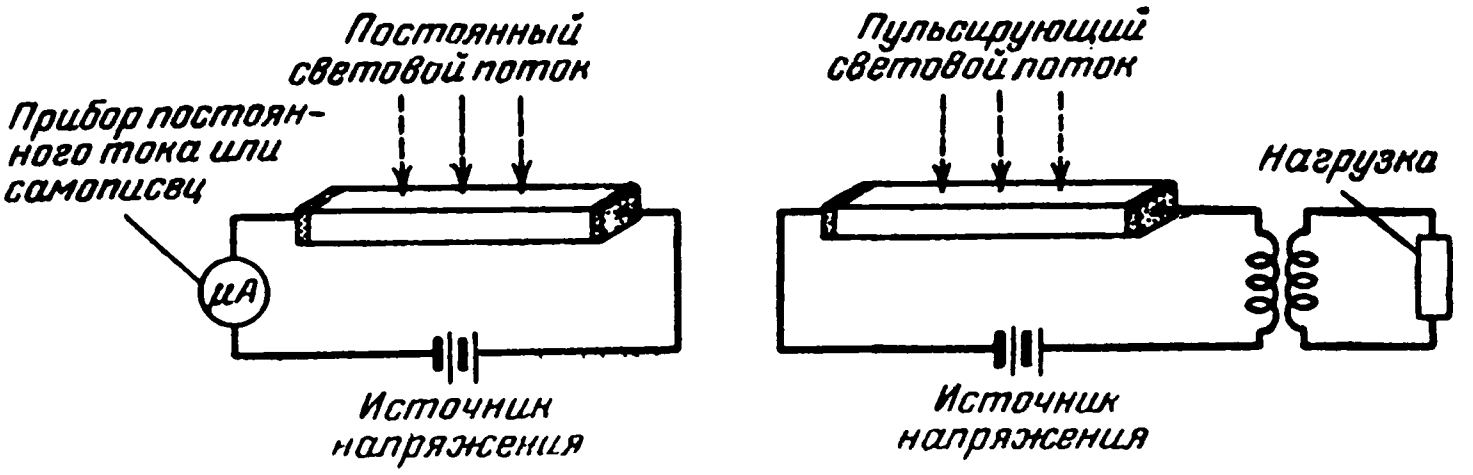


Рис. 8-2. Фотосопротивление может работать в цепи постоянного или переменного тока.

Фотосопротивление всегда используется в цепи с источником напряжения. На рис. 8-2 слева показана цепь постоянного тока, в которой выходной ток фотоэлемента является мерой падающего на него светового потока. Такая цепь может быть использована, например, с самописцем для регистрации интенсивности солнечного света. На том же рисунке справа приведена типичная цепь переменного тока для индикации быстрых колебаний светового потока, подобная используемой для чтения фотоэлектрической записи на звуковой дорожке кинофильма.

8-1-1. Селеновые фотосопротивления

На рис. 8-3 показана схема селенового фотосопротивления [Л. 1]. Оно изготовляется заливкой слоем селена электродных проволочек, выполненных в виде двойной обмотки. Из-за малой толщины этого слоя и высокого удельного сопротивления селена (около 1 000 *ом·см*) приходится принимать меры к тому, чтобы привести полное сопротивление описанной конструкции к практически удобной величине. Технологически прибор выполняется путем намотки двух тонких никелевых или платиновых проволочек на пластинку слюды с расстоянием между ними, равным двум или трем их диаметрам. Затем на пластинку наносится тонкий слой расплавленного селена, пластинка охлаждается и отпускается для перевода селена в полупроводящую аллотропическую форму. В результате этих операций проволочки оказываются полностью залитыми в слой селена толщиной от 0,125 до 0,25 *мм*. Контакты между металлическими проволочками и селеном не выпрямляют. При попадании света на какой-либо участок селенового слоя происходит местное увеличение проводимости между проволочками на освещенном участке. Если освещается вся площадь полупроводника, то этот эффект распространяется на все параллельные участки фотоэлемента.

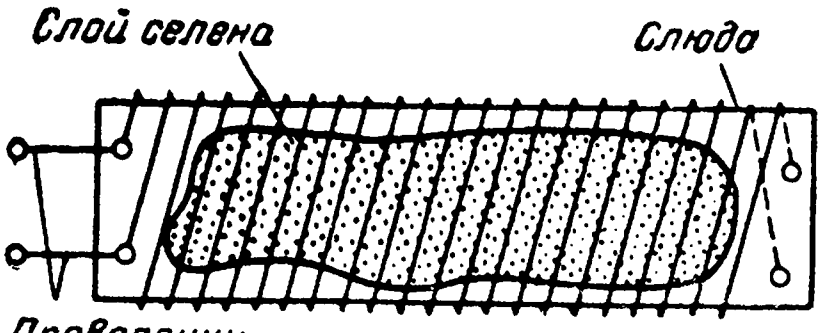


Рис. 8-3. Схема фотосопротивления в виде залитой слоем селена редкой обмотки, навитой двумя тонкими проволочками на изолирующем основании.

В зависимости от диаметра проволочек и расстояния между ними, а также от толщины и режима отжига селеного слоя темновое сопротивление такого фотоэлемента практически удобных размеров колеблется от нескольких десятков тысяч ом до нескольких мегом. Это сопротивление может понижаться до 5% начальной величины при освещении фотоэлемента прямым солнечным светом. Такой фотоэлемент чувствителен ко всем лучам видимого света, но чувствительность его понижается к красному концу спектра, а прозрачность селена повышается с увеличением длины волны света. На рис. 8-4 показана кривая спектральной чувствительности селенового фотосопротивления [Л. 2].

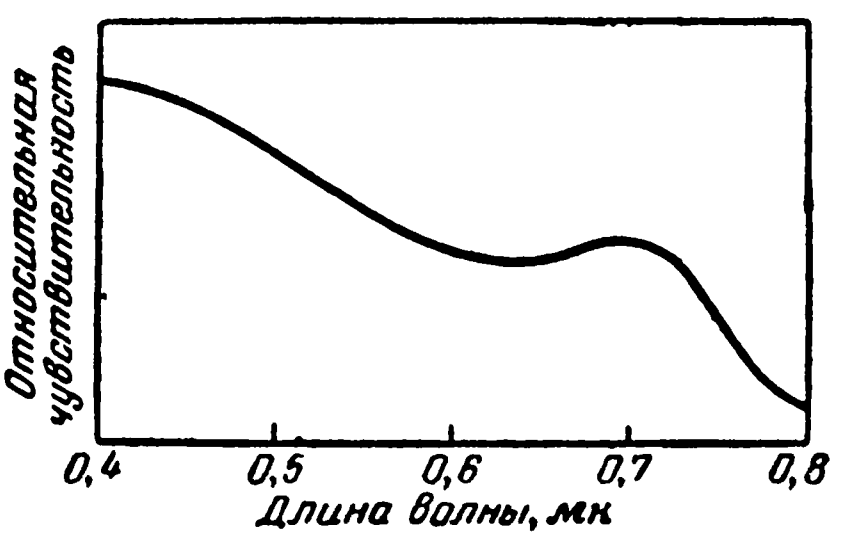


Рис. 8-4. Кривая спектральной чувствительности селенового фотоэлемента охватывает всю видимую часть спектра.

Выполнение электродов в форме проволок является лишь одним из возможных вариантов для селеновых фотоэлементов. Разработан фотоэлемент [Л. 3], в котором тонкий слой селена наносится на керамическую пластинку, предварительно покрытую металлической пленкой. Эта пленка прорезается насквозь до керамического основания извилистой линией, так что металлическое покрытие оказывается состоящим из двух частей, разделенных прорезанной линией. Наносимый затем селеновый слой перекрывает зазор между двумя электрода-

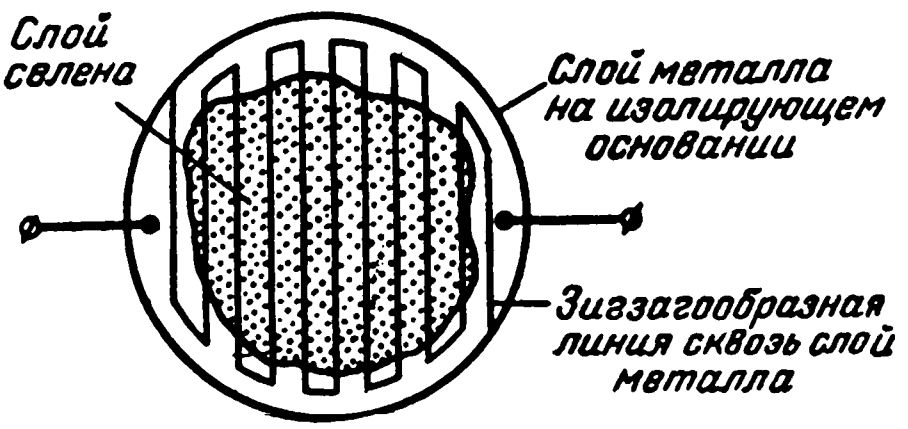


Рис. 8-5 Схема фотосопротивления в виде тонкого слоя селена на металлической подложке, разделенной на две части гравировкой.

ми. На рис. 8-5 показан один из вариантов электродов такого фотоэлемента. В современной практике селеновый слой наносится испарением в вакууме, что позволяет получить более тонкий и однородный слой, чем при намазывании расплавленного селена. После отжига нанесенный селеновый слой превращается в полупроводник, обнаруживающий внутренний фотоэффект. Поскольку свойства селеновых фотоэлементов изменяются под влиянием влаги окружающей среды, в современной производственной практике поверхность селена защищается покрытием из влагозащитного лака или заключением всего элемента в герметическую оболочку.

8-1-2. Серно-свинцовые фотосопротивления

Фотоэлементы из сульфида свинца выполняются в основном подобно селеновым фотоэлементам, т. е. в виде тонкого слоя сульфида свинца, перекрывающего зазор между двумя электродами. Эти электроды могут быть

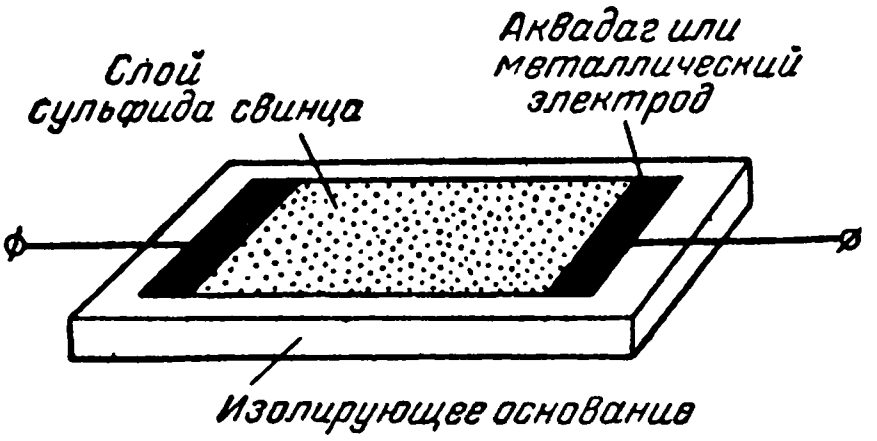


Рис. 8-6. Схема серно-свинцового фотоэлемента, в котором пленка частично окисленного сульфида свинца образует проводящий мостик между двумя электродами.

приготовлены нанесением слоя аквадага или пленки металла, осажденной в вакууме на стеклянную пластинку. Сульфид свинца также осаждается из пара в вакууме на электродную систему, в результате чего получается структура, подобная изображенной на рис. 8-6. После осаждения пластинка отжигается в воздухе или какой-либо другой атмосфере, содержащей кислород. Эта обработка оказывает большое влияние на характеристики фотоэлемента. Цель ее — превращение некоторых участков пленки сульфида свинца в закись свинца, чтобы за счет явления умножения, подобного описанному в гл. 22, повысить светочувствительность слоя сульфида свинца.

Опубликовано много вариантов процесса отжига [Л. 5 и 6].

Спектральная чувствительность сульфида свинца зависит от природы и характера термообработки В лите-

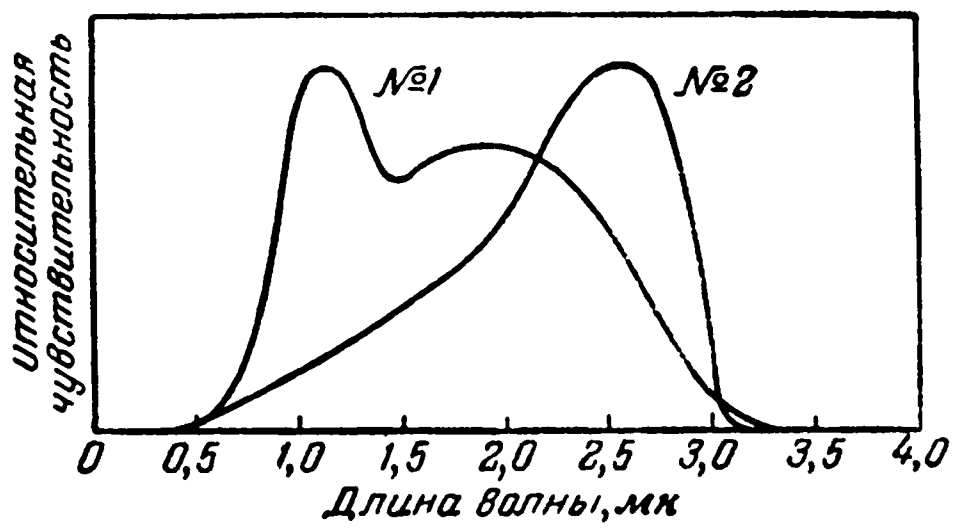


Рис. 8-7. Зависимость спектральной чувствительности серно-свинцового фотоэлемента от режима обработки в процессе его изготовления.

ратуре описаны очень различные формы спектральных характеристик. На рис. 8-7 показаны две крайние характеристики, приведенные к одной и той же макси-

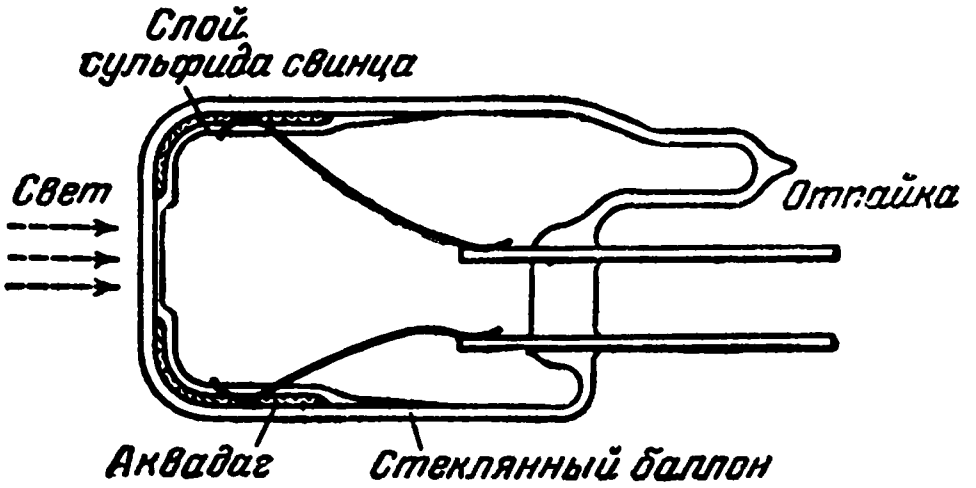


Рис. 8-8. Серно-свинцовый фотоэлемент в вакуумном стеклянном баллоне.

мальной величине, полученные при разных режимах термообработки.

Очувствленные к свету пленки из смеси сульфида свинца и закиси свинца неустойчивы по отношению к атмосферному воздуху. Поэтому серно-свинцовые фотоэлементы изготовляются так, чтобы светочувствительный слой после нанесения и термообработки мог быть запаян в вакуумном баллоне. На рис. 8-8 показан схематически разрез вакуумного серно-свинцового фотоэлемента.

При изготовлении его на внутренние стенки стеклянного баллона наносятся аквадаговые электроды желательной формы и с требуемым зазором. Затем наносится слой сульфида свинца и термически обрабатывается, после чего баллон откачивается и отпаивается. Промышленные серно-свинцовые фотоэлементы выпускаются с темновым сопротивлением в пределах от 100 000 *ом* до 10 *Мом*.

8-1-3. Германиевые точечные фотосопротивления

На рис. 8-9 показано фотосопротивление с необычной геометрией электродов [Л. 7]. Полупроводник (германий) имеет форму шлифованной шайбы, на одной стороне которой вытравлена лунка, так что толщина

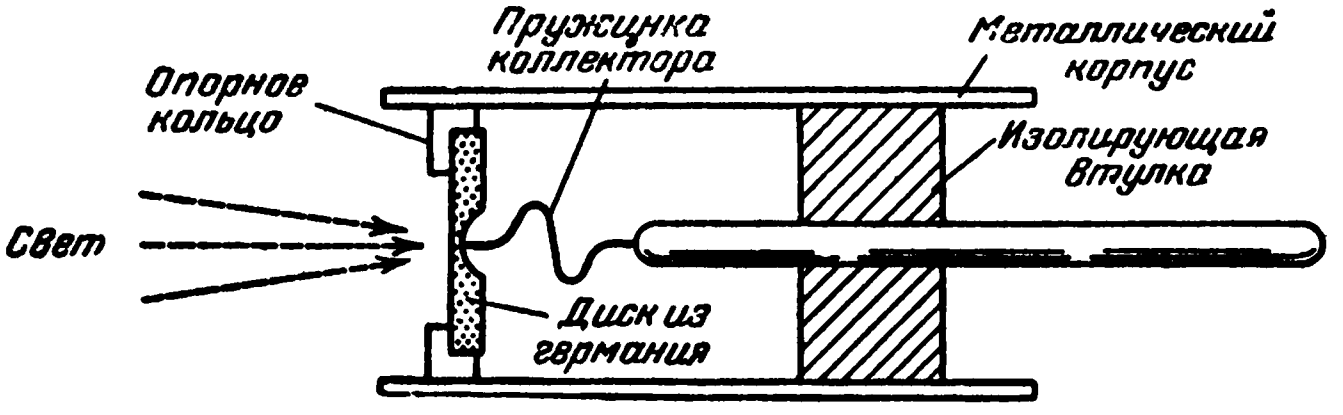


Рис. 8-9. Схематический продольный разрез точечного германиевого фототранзистора.

шайбы в центре составляет около 0,075 *мм*. В самой тонкой части с шайбой соприкасается точечный электрод, называемый коллектором, который образует выпрямляющий контакт; контакт шайбы с держателем по ее окружности является невыпрямляющим. Вся конструкция собирается в цилиндрическом металлическом корпусе. Свет, который может быть сконцентрирован линзой, направляется на небольшой участок полупроводника непосредственно против коллектора. Такой фотоэлемент подобен двустороннему коаксиальному транзистору (описанному в следующей главе) во всех отношениях, за исключением того, что возбуждающий электрический заряд, вместо того чтобы вводиться в полупроводник с другого электрода, освобождается в полупроводнике под действием света. По этой причине описываемый прибор назван фототранзистором [Л. 8]. На рис. 8-10 приведены фотографии одного из типов фототранзисторов.

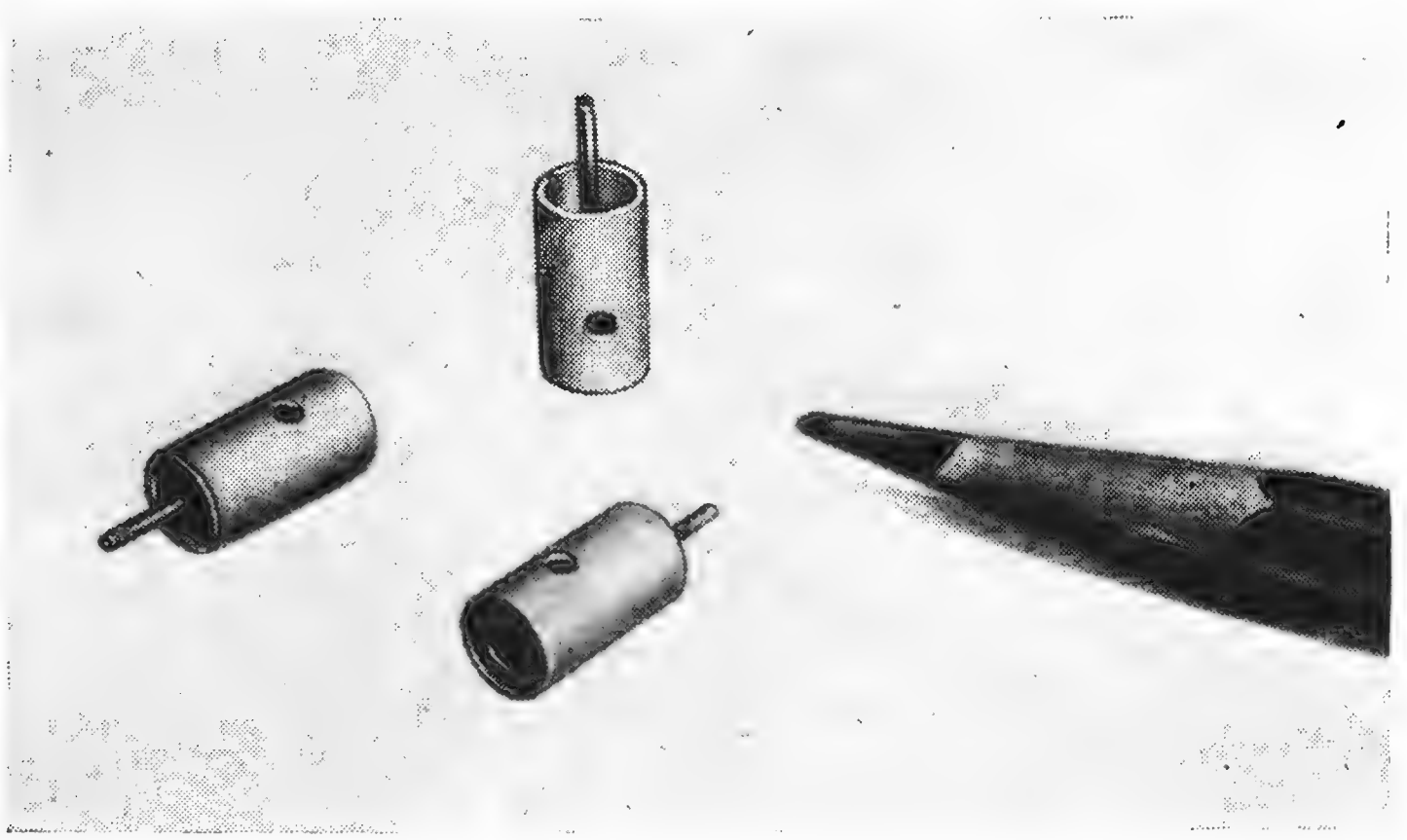


Рис. 8-10. Размеры экспериментальных образцов точечных фототранзисторов.

Обычно фототранзистор находится под обратным напряжением (если рассматривать его как германиевый диод). На рис. 8-11 приведено семейство электрических характеристик фототранзистора при различных световых потоках. Если использован германий *n*-типа и соблюдены указанные направления напряжения и тока, то наиболее интересные части этого семейства кривых располагаются в третьем квадранте, как показано на графике. «Темновая» кривая представляет собой характеристику обратного тока германиевого диода очень плохого качества. Другие кривые показывают, как характеристика смещается к бóльшим токам при увеличении светового потока, падающего на активный участок фотоэлемента. Характерно, что кривые расположены на неравных расстояниях друг от друга, сближаясь при более высоких световых потоках. Эта нелинейность — специфическая особенность всех фотосопротивлений. Она обусловлена тем, что при высоких освещенностях концентрация освобожденных светом дырок и электронов выше, чем при низких уровнях освещенности, и что быстрота рекомбинации при этом оказывается соответственно выше, так что продолжительность существования освобожденных частиц получается меньшей. Следовательно, выходной ток такого фотоэлемента более или

менее точно следует зависимости вида $I \sim L^n$, где L — световой поток и n — показатель, величина которого находится в пределах от 0,5 до 1.

Спектральная характеристика германиевого фототранзистора, находящегося под рабочим напряжением, приведена на рис. 8-12, на которой чувствительность выражена через квантовый выход. Квантовый выход пред-

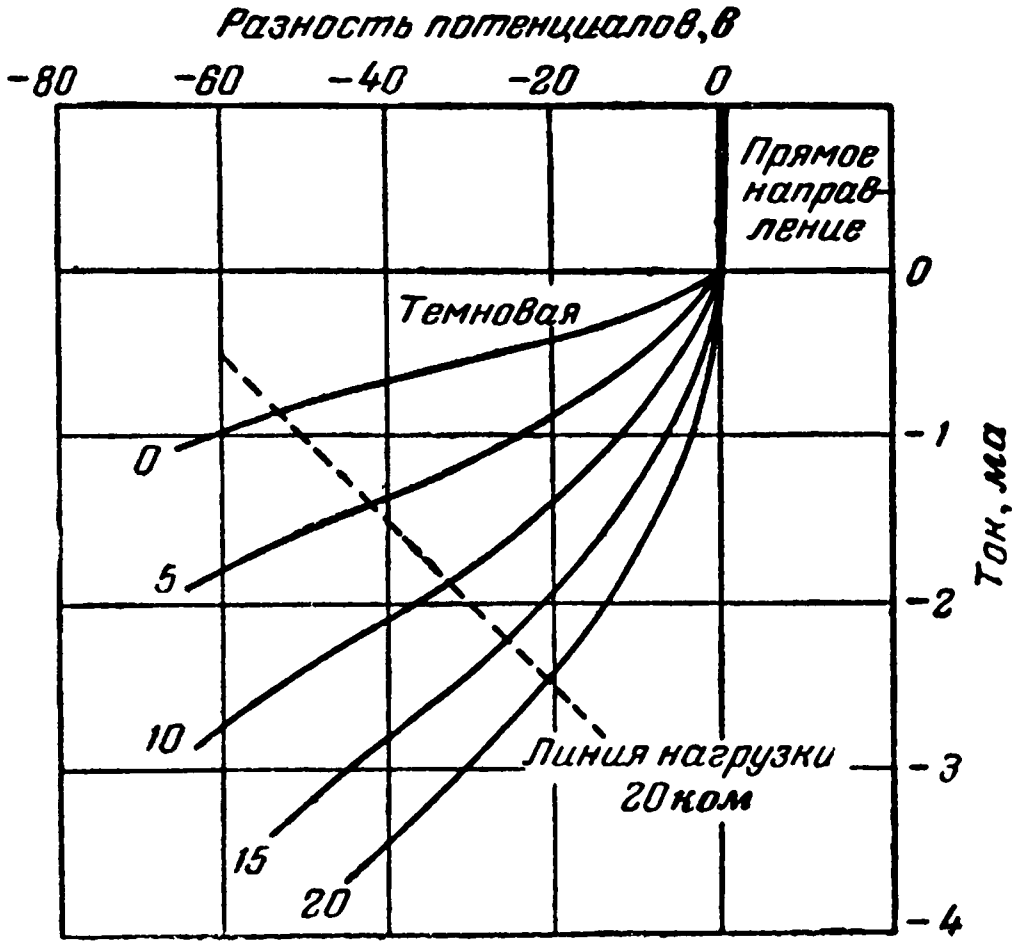


Рис. 8-11. Зависимость электрических характеристик точечного германиевого фототранзистора от падающего светового потока. Цифры у кривых указывают величину светового потока в миллилюменах. Свет — от лампы накаливания с вольфрамовой нитью при цветовой температуре 2 870° К.

ставляет собой количественный критерий, часто используемый для характеристики работы фотоэлемента. Эта величина определяется как отношение числа электронов фотоэлектронного тока во внешней цепи к числу фотонов [1], падающих на активную поверхность фотоэлемента. У фототранзисторов наблюдался выход, достигающий трех или четырех электронов на фотон, что указывает на существование некоторого вторичного процесса умножения зарядов, увеличивающего первичный заряд, фактически освобождаемый при фотоэффекте. Нужно

[1] Фотонная теория света рассмотрена в гл. 22.

отметить, что германиевый фототранзистор, обнаруживает наибольшую чувствительность при длине волны света примерно 1,4 *мк* в инфракрасной области, далеко за красной границей видимой части спектра; но зато на этом участке спектра он особенно подходит для использования со многими искусственными источниками света.

Точечному фототранзистору свойственна та полезная особенность, что площадь наиболее чувствительного уча-

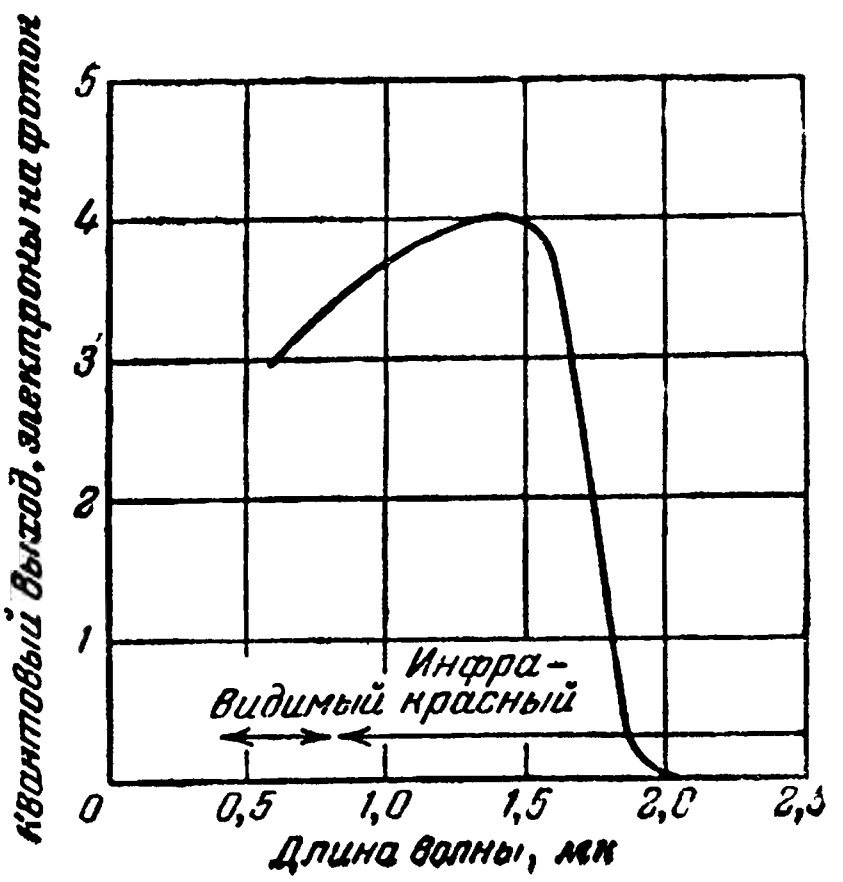


Рис. 8-12. Кривая спектральной чувствительности германиевого фототранзистора охватывает как видимый, так и инфракрасный участки спектра.

стка на освещенной стороне германиевой шайбы в точке, лежащей непосредственно против коллекторного контакта, составляет всего несколько сотых квадратного миллиметра. На рис. 8-13 показана кривая распределения чувствительности фототранзистора, полученная путем передвижения небольшого пятна света по передней поверхности фототранзистора по диаметру германиевой шайбы. Наиболее чувствительный участок фотоэлемента имеет обычно около 0,25 *мм* в диаметре. Свет, падающий вне этого участка, вызывает лишь незначительный электрический эффект во внешней цепи. Кривые на рис. 8-11 сняты в условиях, когда весь световой поток (принятый в качестве параметра для этих кривых) сфокусирован целиком в пределах указанного участка. Подобным же образом были определены квантовые выхо-

ды, приведенные на рис. 8-12. Малый размер чувствительной части фототранзистора позволяет рекомендовать его для применения в тех случаях, когда желательна высокая разрешающая способность, как, например, в случае передачи изображений. Еще одним достоинст-

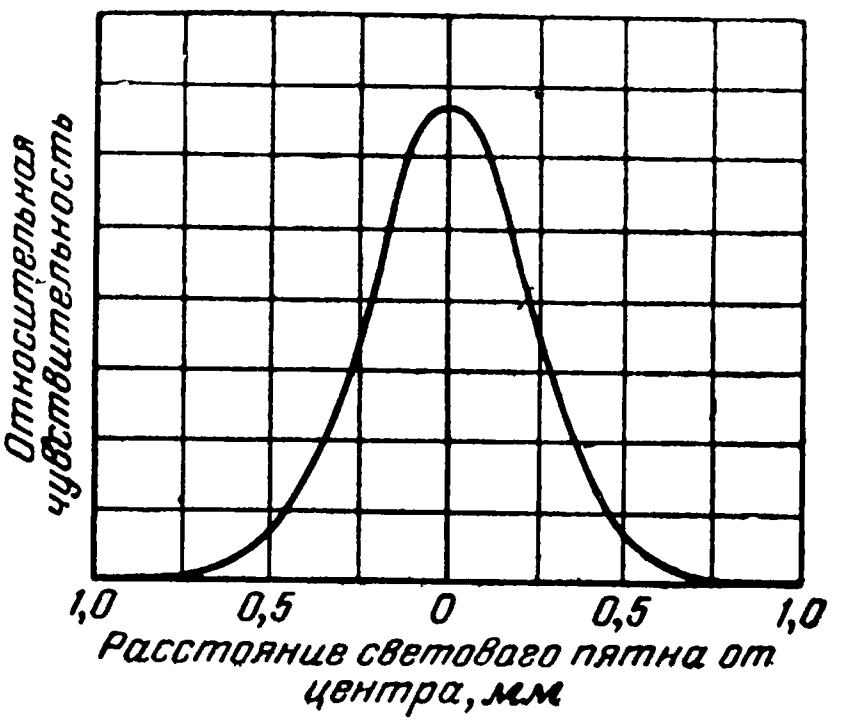


Рис. 8-13. Максимальная чувствительность германиевого фототранзистора ограничена небольшой площадью в центре германиевого диска

вом германиевого фототранзистора, не всегда свойственным другим типам полупроводниковых фотоэлементов, является быстрота реакции на изменение интенсивности светового потока. Экспериментально показано, что его частотная характеристика чувствительности остается плоской по крайней мере до 200 *кгц*.

8-1-4. Фотоэлементы с *p-n* переходом

В то время как связь фотогальванических явлений с выпрямительными контактами известна с 1883 г. [Л. 9], связь фотопроводимости с выпрямительными *p-n* переходами выяснена была сравнительно недавно. В 1941 г. было обнаружено [Л. 10], что на электрическую характеристику естественного *p-n* перехода в кремнии влияет освещение. В 1947 г. появилось [Л. 11] сообщение о внутреннем фотоэффекте, обнаруживаемом на выпрямительном точечном контакте, расположенном последовательно с естественным *p-n* переходом в германии. В 1951 г. были описаны [Л. 12] свойства германиевого

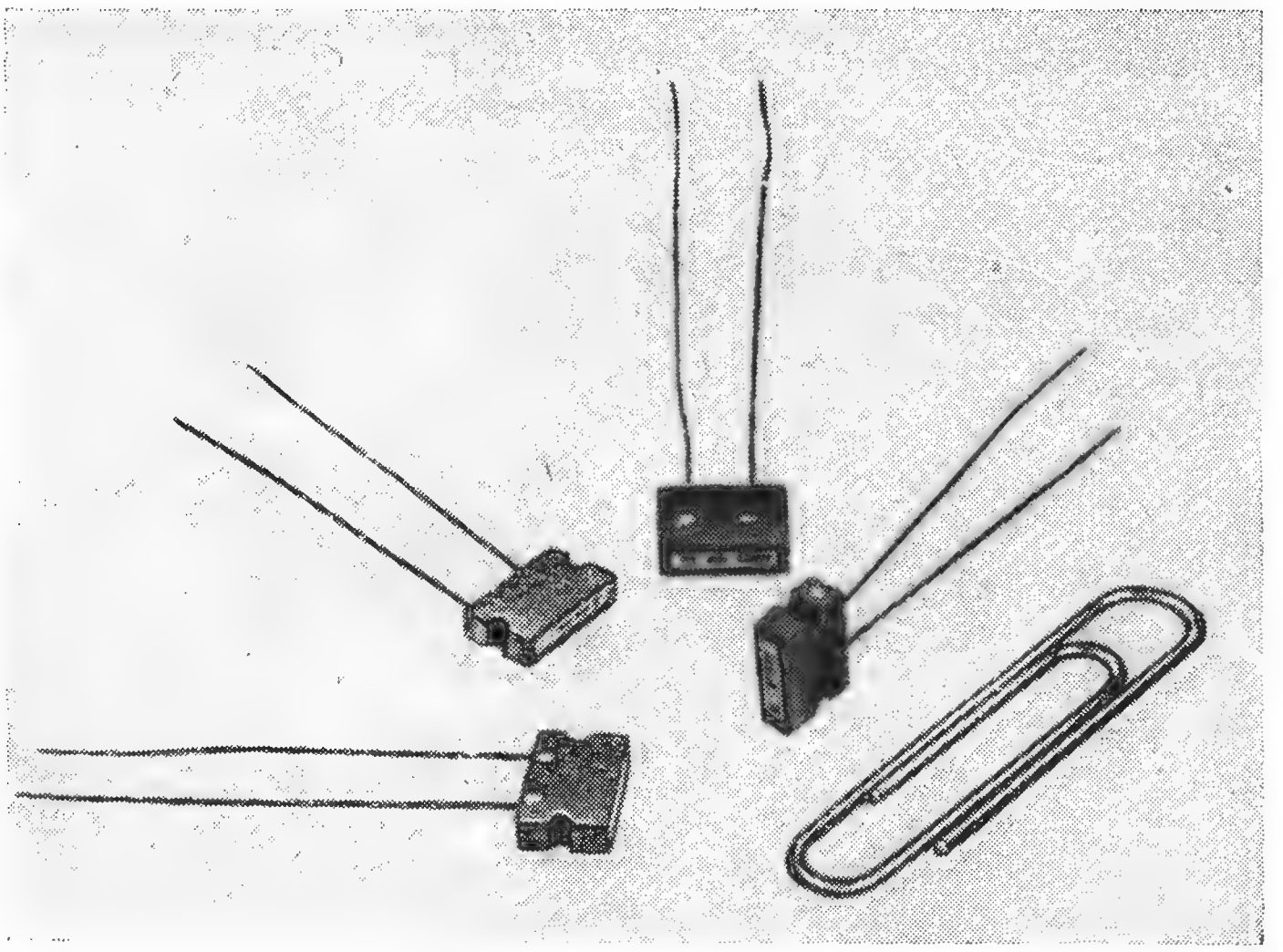


Рис. 8-14 Внешний вид фотоэлементов с *p-n* переходом, заделанных в пластмассовые оболочки.

фотоэлемента с выращенным *p-n* переходом, подобного выпускаемым в настоящее время.

Изготовление фотоэлементов с *p-n* переходом производится в точности так же, как и диодов с *p-n* переходом (см. гл. 5), за исключением лишь окончательной заделки в непрозрачную оболочку. В современных фотоэлементах с *p-n* переходом полупроводниковый элемент полностью заливается в корпус из прозрачной пластмассы, по крайней мере одна сторона которого выполняется в виде плоской или выпуклой фокусирующей поверхности. Если элемент включен в цепь постоянного напряжения такой полярности, что фотоэлемент, рассматриваемый как диод, находится под обратным напряжением, то в темноте через цепь протекает лишь небольшой обратный ток насыщения. Освещение фотоэлемента приводит к появлению добавочного тока, величина которого почти точно пропорциональна световому потоку в широком диапазоне вплоть до получаемого при прямом солнечном освещении.

На рис. 8-14 приведены фотографии некоторых типов фотоэлементов с *p-n* переходами. На рис. 8-15 дан

разрез фотоэлемента, показывающий расположение элемента с *p-n* переходом в пластмассовой оболочке. Семейство электрических характеристик этого фотоэлемента показано на рис. 8-16 [Л. 13]. Темновая кривая соответствует темновому току всего в несколько микроампер и, как и следует ожидать для обратной характеристики выпрямителя с *p-n* переходом, обнаруживает хорошее насыщение тока при напряжениях в пределах 1 *в* и выше. Как в темноте, так и при освещении в широких пределах обратных напряжений поддерживаются почти постоянные величины обратных токов насыщения.

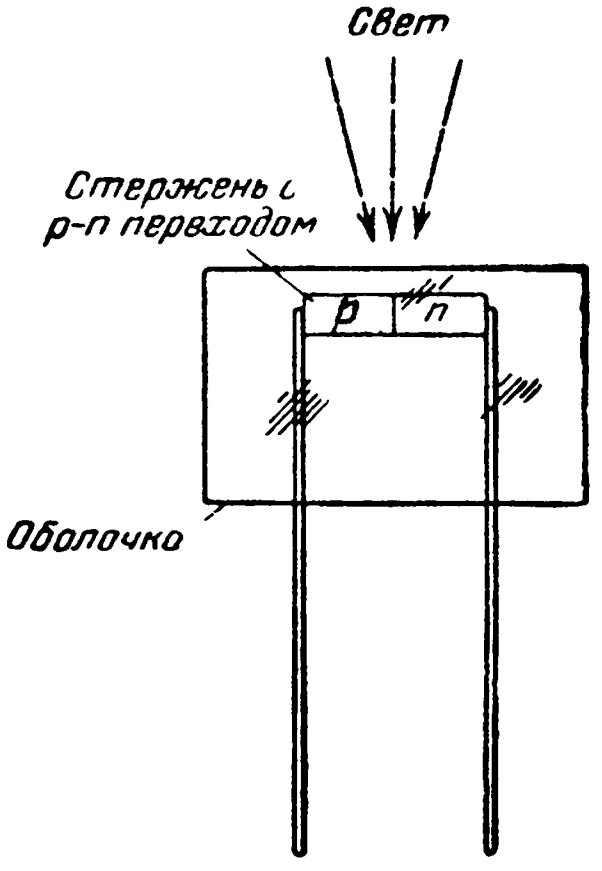


Рис. 8-15. Устройство фотоэлемента с *p-n* переходом в пластмассовой оболочке.

Эта особенность — почти полная независимость тока от приложенного напряжения — позволяет предложить для фотоэлемента с *p-n* переходом эквивалентную схему (рис. 8-17), в которой активным элементом является генератор не напряжения, а тока, обеспечивающий ток kL, пропорциональный световому потоку, где L — световой поток, а k — коэффициент пропорциональности. Схема на рис. 8-17 напоминает выпрямитель с генератором тока, включенным параллельно элементам запирающего слоя. Для германиевого элемента с *p-n* переходом, характеристики которого даны на рис. 8-17, при напряжении питания 50 *в* r может иметь величину 100 *ом*, R—50 *Мом*, и C—10 *пф*. I_0 может составлять 1—2 *мка* и k—около 30 *ма/лм*. Если внешнее сопротивление нагрузки несравнимо по величине с R, то ток нагрузки почти не зависит от напряжения питания и величины сопротивления нагрузки.

Замечательная особенность характеристик на рис. 8-16 заключается в том, что из-за малой величины темнового тока отношение тока при освещении к темновому току для данного светового потока больше, чем у любого другого типа фотосопротивления, даже несмотря на то, что квантовый выход для фотоэлемента с *p-n*

переходом никогда не бывает выше одного электрона на фотон. Эти приборы часто предпочитаются другим, таким, как точечные фототранзисторы или серно-свинцовые фотоэлементы, вследствие большой величины отношения тока при освещении к темновому. В тех случаях, когда фотоэлемент используется в цепи постоянного

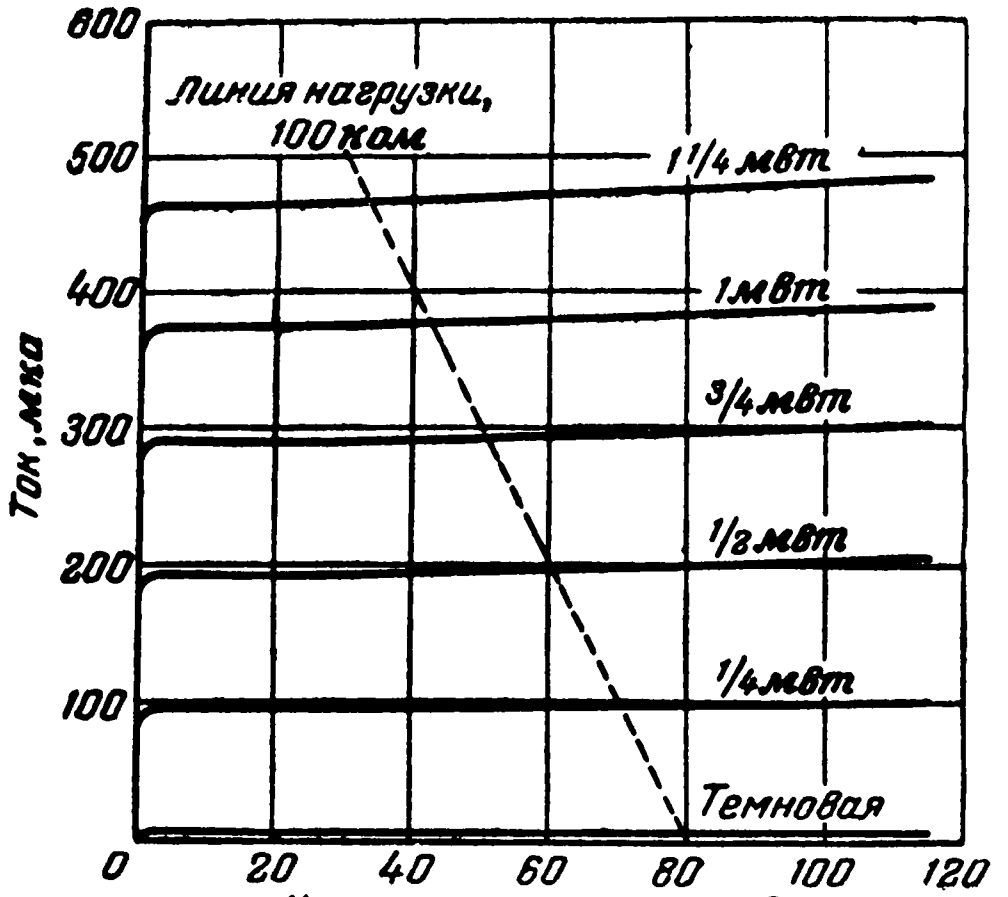


Рис. 8-16 Характеристики германиевого фототранзистора с *p-n* переходом при освещении области в пределах ±0,125 *мм* от перехода. Величины световых потоков даны в милливаттах полной мощности, излучаемой лампой накаливания с вольфрамовой нитью при цветовой температуре 2 870 К.

тока для индикации наличия или отсутствия освещения, это отношение важнее, чем прирост тока при освещении.

Подобно точечному фототранзистору фотоэлемент с *p-n* переходом чувствителен только к световому потоку, падающему на переход или очень близко к нему. Причину этого ограничения чувствительной поверхности можно уяснить с помощью рис. 8-18, на нижней части которого показано распределение потенциала в полупроводниковом элементе, снятое при перемещении зонда по поверхности стержня от одного конца до другого, пере-

секая на пути *p-n* переход, находящийся под обратным напряжением. Почти все напряжение, приложенное к элементу от источника, падает на высоком сопротивлении запирающего слоя. Именно в этом слое, и только здесь, существует сильное электрическое поле. Ток во внешней цепи проходит только тогда, когда заряженная частица пересекает переход. Если в толще полупроводника в его *n*- или *p*-участке в области, удаленной от перехода, освобождаются под действием света электроны или дырки, то местное поле оказывается слишком слабым, чтобы разделить и собрать эти заряды. После пребывания в течение нескольких микросекунд по соседству с местом их образования эти частицы исчезают из-за рекомбинации, не внося заметного прироста в общий ток цепи. Наоборот, заряды, освобождаемые светом, даже необязательно на самом переходе, но поблизости от

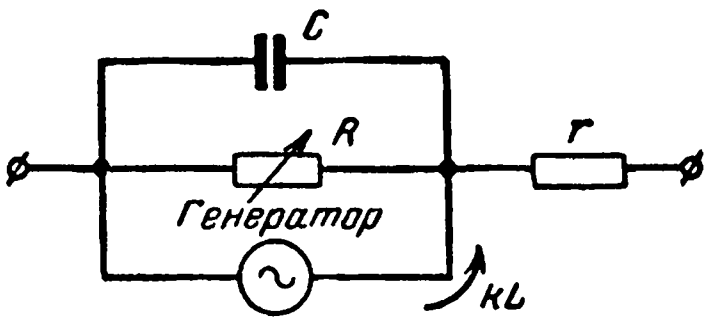


Рис 8-17. Эквивалентная схема фотосопротивления с переходом Генератор тока включен параллельно переходу.

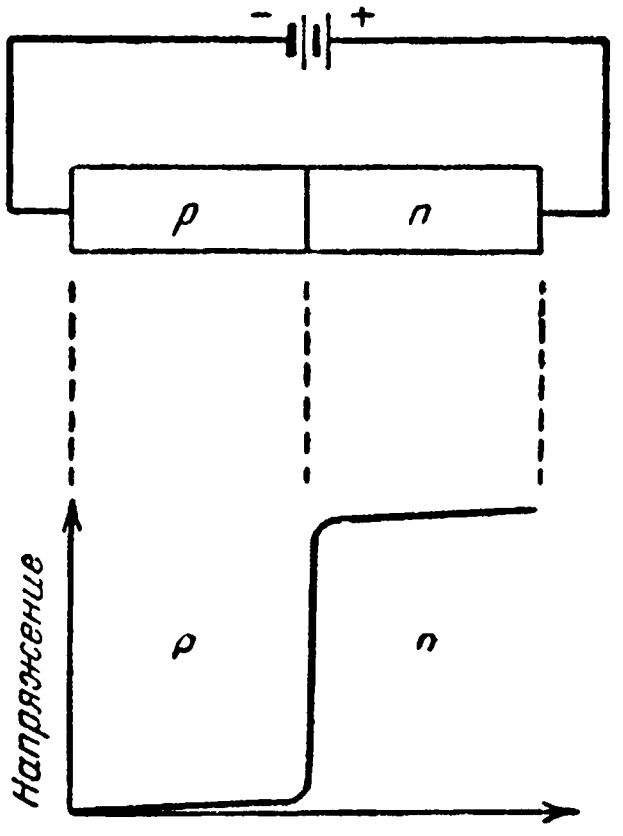


Рис 8-18. В фотосопротивлении с *p-n* переходом бóльшая часть напряжения питания падает на переходе. Электрическое поле в остальных частях элемента пренебрежимо мало.

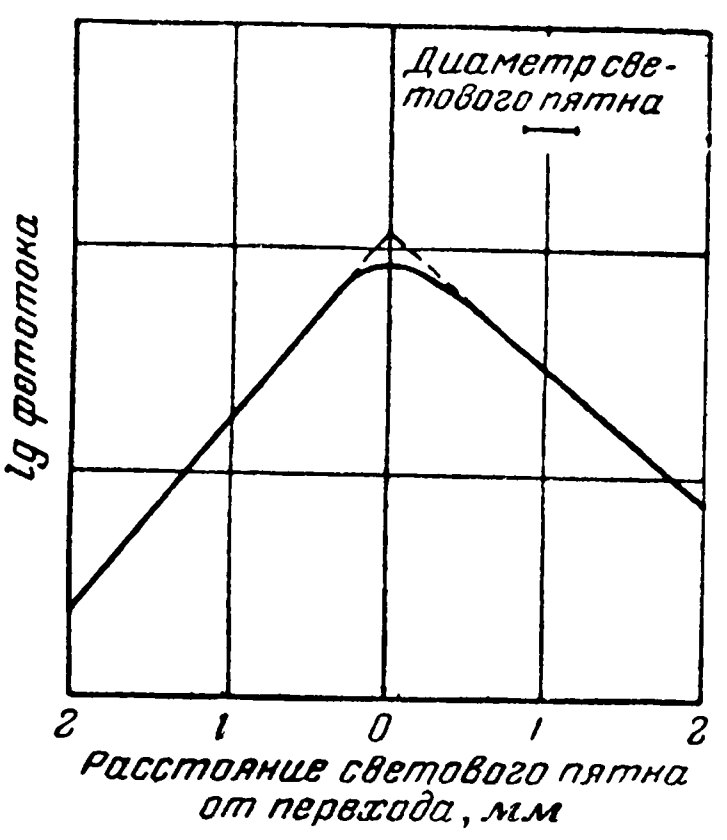


Рис. 8-19. Экспоненциальная зависимость фототока от расстояния между световым пятном и переходом. Пятно перемещается через линию *p-n* перехода, находящегося под обратным напряжением.

него, могут успеть до рекомбинации продиффундировать в область перехода, где они разделяются и собираются полем запирающего слоя. Вероятность этого понижается экспоненциально с увеличением расстояний от перехода до места их образования [Л. 14]. Этот факт, так же как и то обстоятельство, что переход не бесконечно тонок и что границы электрического поля получаются несколько размытыми на краях, приводит к распределению чувствительности по длине элемента, показанному на рис. 8-19. Измерения, проделанные на германиевых фотоэлементах с *p-n* переходом с помощью световых вспышек с очень коротким временем нарастания и спада, показали, что задержка реакции фотоэлемента на изменение светового потока порядка микросекунды, если свет падает точно на область перехода, и нескольких микросекунд, если бóльшая площадь фотоэлемента освещается рассеянным световым потоком. В последнем случае задержка обусловливается временем диффузии частиц, образовавшихся на некотором расстоянии от перехода.

Поскольку основы процесса освобождения зарядов во всех германиевых фотоэлементах одни и те же, все они имеют в основном одинаковое спектральное распределение чувствительности по длинам волн падающего светового потока. Это распределение для германия показано на рис. 8-12.

8-1-5. Фотосопротивления со многими переходами

В результате усиленной разработки транзисторов появился еще один вид фотоэлемента со структурой типа *n-p-n* [Л. 15]. Схема такого фотоэлемента из германия или кремния показана на рис. 8-20; он состоит из тонкого слоя материала *p*-типа, заключенного между двумя электродами из материала *n*-типа. Весь элемент вырезается из монокристаллической заготовки, приготовленной путем контролированного введения примесей в процессе ее выращивания.

Электрические характеристики германиевого фотоэлемента типа *n-p-n* для различных освещенностей даны на рис. 8-21. Характерно, что при сравнимых световых потоках получаются фототоки примерно в 25 раз выше, чем для простого фотоэлемента с одним *p-n* переходом. Это означает, что квантовый выход для фо-

тоэлемента с двойным переходом составляет примерно 25 электронов на фотон. Так как квантовый выход первичного фотоэффекта, в результате которого при поглощении одного фотона света образуются один электрон и одна положительная дырка, никогда не бывает выше единицы, то такой высокий квантовый выход, какой наблюдается в фотоэлементе с двойным переходом, указывает на существование вторичного процесса умножения. Это действительно имеет место. Физические причины этого умножения те же, которые вызывают эффект умножения тока в транзисторах. Основы эффекта такого умножения разобраны в гл. 22.

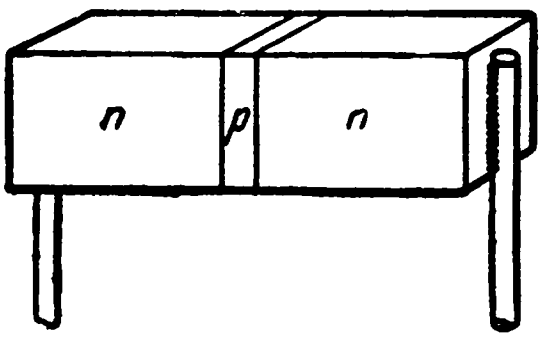


Рис. 8-20. Структура фотоумножительного сопротивления типа *n-p-n* с двумя близко расположенными *p-n* переходами.

Заметим также, что характеристики фотоэлемента типа *n-p-n*, представленные на рис. 8-21, после выхода из области малых напряжений, где они обнаруживают тенденцию к насыщению тока, перегибаются и об-

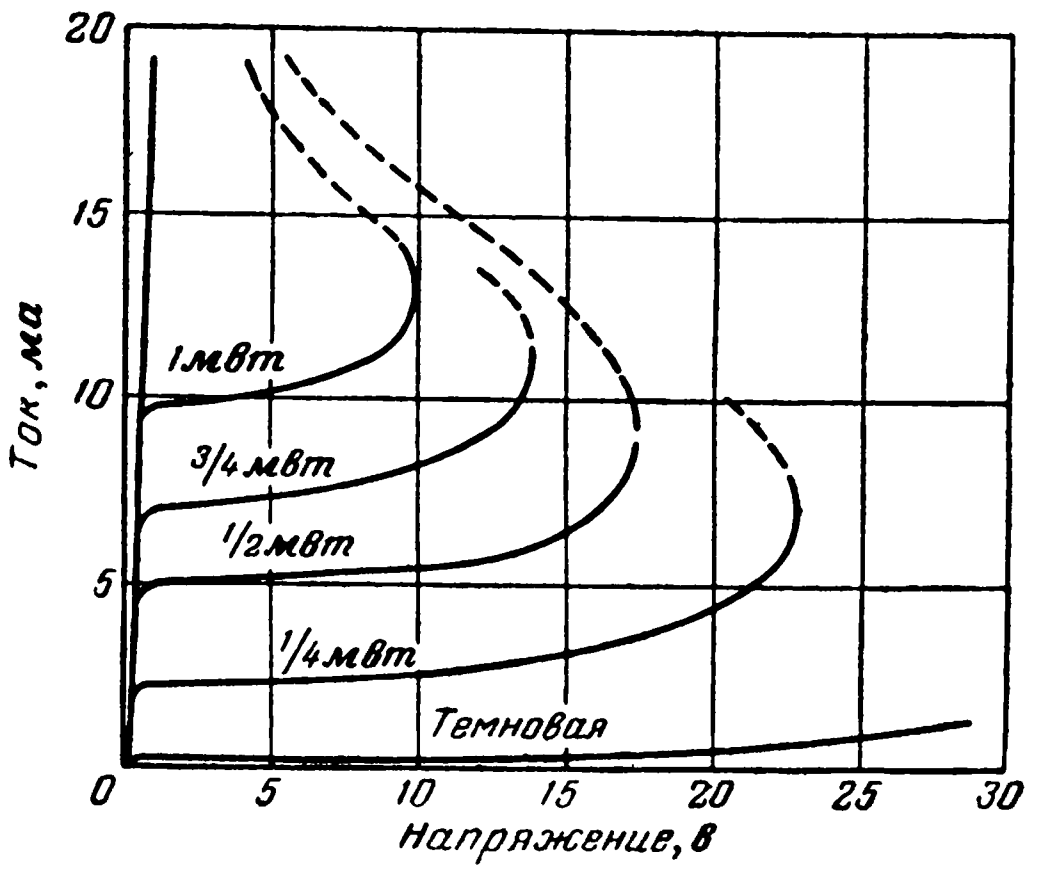


Рис. 8-21. При равных входных мощчостях излучения такое фотоумножительное сопротивление типа *n-p-n* дает примерно в 25 раз больший фототок, чем фотосопротивление с *p-n* переходом (рис. 8-16).

разуют участки с отрицательными сопротивлениями переменному току, причем величина максимального напряжения для каждой кривой зависит от светового потока, при котором она снималась. Это поведение, несомненно, вызвано максимумом термистерного эффекта, имеющего место в полупроводнике вследствие нагрева током. Цепь, состоящая из *n-p-n* фотоэлемента, батареи и нагрузочного сопротивления, находящаяся в темноте (малый ток), может быть переведена, как триггер, в состояние с большим током через пик напряжения в результате освещения.

Считается, что особенности характеристик серно-свинцовых, серно-таллиевых и других полупроводников фотоэлементов, обнаруживающих результирующий квантовый выход, больший, чем один электрон на фотон, обусловлены структурой полупроводникового слоя. Так, например, считают, что в процессе очувствления при изготовлении серно-свинцовых фотоэлементов после нанесения слоя сульфида свинца происходит частичное восстановление микроскопических островков и участков слоя до закиси свинца, которая отличается по типу проводимости от сульфида. Таким образом, слой оказывается пронизанным сетью двойных переходов *n-p-n*, в которых происходит умножение тока, точно так же, как и в отдельном двойном переходе типа *n-p-n*, описанном выше.

8-2. ФОТОГАЛЬВАНИЧЕСКИЕ ЭЛЕМЕНТЫ

Способность выпрямляющего контакта или *p-n* перехода генерировать э. д. с. при освещении позволила создать приборы, называемые фотогальваническими элементами. Эти элементы генерируют свою собственную э. д. с. и используются в схемах без какого-либо внешнего источника питания. Фотогальванический эффект наблюдается при освещении края шайбы меднозакисного выпрямителя или края металлического контактного слоя селенового выпрямителя. Фото-э. д. с. возникает в кристаллических диодах при освещении точки контакта, и фотоэлемент с *p-n* переходом обладает наряду с фотопроводимостью фотогальваническими свойствами. Однако при конструировании специальных фотогальванических элементов изыскивается такая геометрия электродов, которая обеспечивает более легкий доступ света

к выпрямительному контакту или переходу, чем у выпрямителей обычной конструкции. В следующих параграфах будут рассмотрены различные варианты устройства контактов, отвечающие указанным требованиям.

Эффективность преобразования света в электрический выходной сигнал у фотогальванических элементов меньше, чем у фотосопротивлений. Это обусловлено тем, что даже при наиболее удачном применяемом в настоящее время расположении электродов, неизбежна потеря за счет отражения и поглощения в различных частях элемента, помимо выпрямительного слоя, значительной части входного светового потока. Кроме того, сам принцип работы прибора исключает возможность использования умножения тока, повышающего квантовый выход некоторых фотосопротивлений. Другим недостатком фотогальванических элементов, который исключает во многих случаях возможность их применения, является то, что наибольшая э. д. с., генерируемая одним элементом, при самых высоких практически используемых световых потоках редко превышает 0,4 или 0,5 *в*. Несмотря на это, фотогальванические элементы пользуются популярностью, так как удобство применения их без внешнего источника питания оказывается более существенным, чем трудности, обусловленные обоими указанными выше недостатками.

Эквивалентная схема фотогальванического элемента с запирающим слоем подобна схеме фотосопротивления, показанной на рис. 8-17, за исключением того, что величина коэффициента пропорциональности k получается меньшей по указанным выше причинам. При увеличении светового потока ток, генерируемый за счет фотоэффекта, также увеличивается, распределяясь между сопротивлением запирающего слоя R и сопротивлением внешней цепи. Поэтому при увеличении светового потока выходное напряжение фотоэлемента повышается. Однако полярность возникающей э. д. с. такова, что она прилагается к *p-n* переходу в *прямом* направлении и, таким образом, с увеличением светового потока R уменьшается и все бóльшая и бóльшая часть фототока теряется за счет этого внутреннего шунтирования. Поэтому ток во внешней нагрузке имеет тенденцию к насыщению при повышении светового потока. На рис. 8-22 дано семейство кривых зависимости напряжения от освещенно-

сти для фотогальванического элемента с *p-n* переходом при различных сопротивлениях внешней цепи. Характеристики фотоэлемента оказываются линейными только при малых световых потоках, но пределы линейности могут быть расширены уменьшением сопротивле-

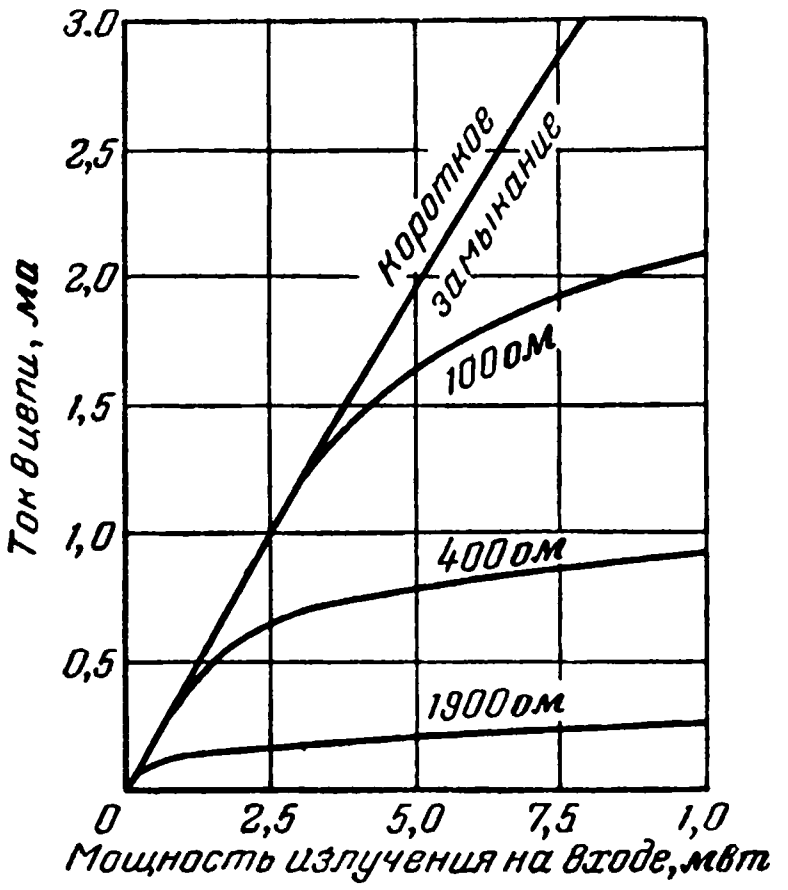


Рис. 8-22. Световые характеристики фотогальванического элемента с *p-n* переходом при разных значениях внешней нагрузки (цветовая температура 2 870° К).

ния нагрузки. На практике фотогальванические элементы часто используются с сопротивлениями нагрузки, величины которых составляют от $^1/_{100}$ до $^1/_{10}$ величины, соответствующей оптимальному согласованию сопротивлений по мощности.

Как и в случае выпрямителей, максимальная частота изменения светового потока, на которую фотоэлемент способен реагировать, ограничена емкостью перехода.

8-2-1. Селеновые фотогальванические элементы

Фотогальванический эффект в контакте между селеном и металлом был впервые обнаружен [Л. 16] в 1877 г. Вскоре после этого [Л. 17] был изготовлен фотогальванический элемент путем нанесения полупрозрачной пленки золота на слой селена, расположенный на медной пластинке (рис. 8-23). Свет, проникающий сквозь золо-

той листок на контакт между золотом и селеном, генерировал э. д. с., измеряемую между золотым электродом и медной пластинкой. Эта ранняя работа не была использована в течение полувека. Только после вторичного открытия фотогальванического элемента [Л. 18] в 1931 г. эти электронные приборы стали хорошо известны. Элементы современного селенового фотогальванического элемента расположены так же, как это было сделано в раннем приборе, но при изготовлении исполь-

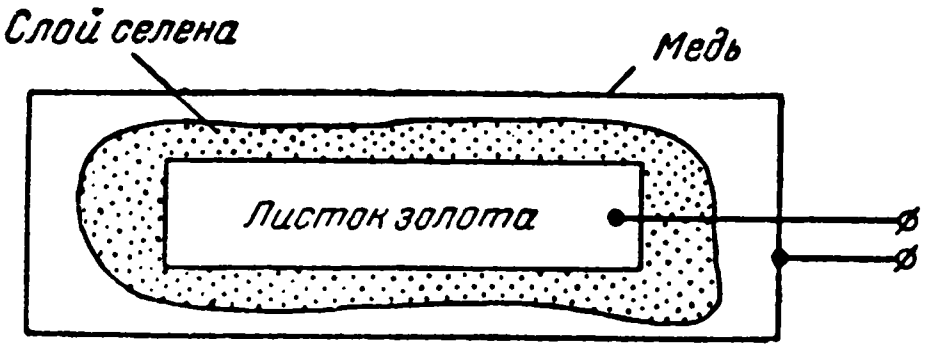


Рис. 8-23 Селеновый фотогальванический элемент. на медную пластину —основание— нанесен слой селена, поверх которого наложен электрод в виде золотого листка.

зуется более современная технология. Подготовка нижнего электрода и селенового слоя производится, так же, как и при изготовлении обычных селеновых выпрямителей, вплоть до момента нанесения внешнего контакта, который у фотоэлементов представляет собой тонкий полупрозрачный слой золота, серебра, алюминия или платины, напыленной разбрызгиванием или испарением в вакууме. Для контакта с этой тонкой пленкой по периферии наносится более толстый слой металла, к которому прижимается контакт в виде металлического кольца. На рис. 8-24 показан схематический разрез такой конструкции. Для защиты от механических повреждений и разрушения влагой весь элемент обычно герметически заделывается в защитный кожух со стеклянным окошечком для света.

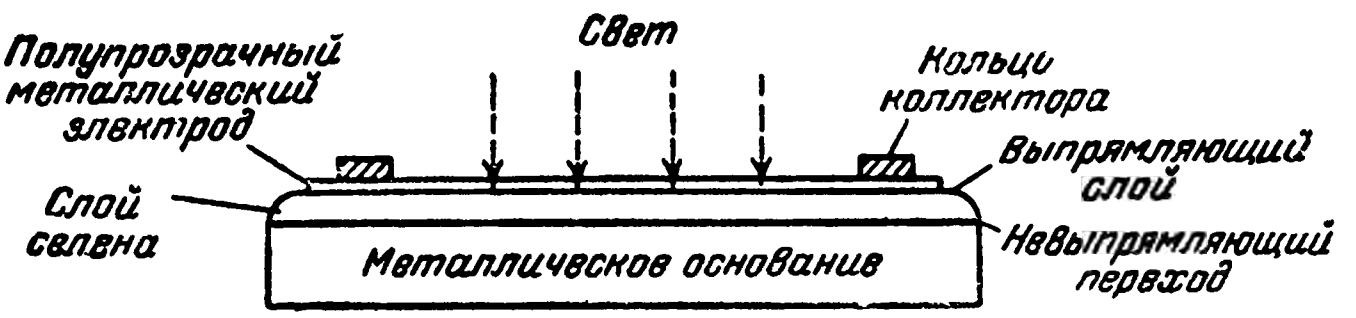


Рис 8-24 Устройство современного селенового фотогальванического элемента.

Промышленные фотоэлементы имеют высокую чувствительность к синему и фиолетовому участкам спектра. Такие фотоэлементы обладают внутренним сопротивлением в несколько тысяч ом и чувствительностью в режиме короткого замыкания в несколько десятых миллиампера на люмен. Фотогальванические приборы находят широкое промышленное применение в переносных фотографических экспонометрах.

8-2-2. Меднозакисные фотогальванические элементы

Вскоре после изобретения меднозакисного выпрямителя был обнаружен [Л. 19] фотогальванический эффект в меднозакисном запирающем слое. Первый фотогальванический элемент представлял собой оксидированную медную пластинку, на которую наложен внешний контакт

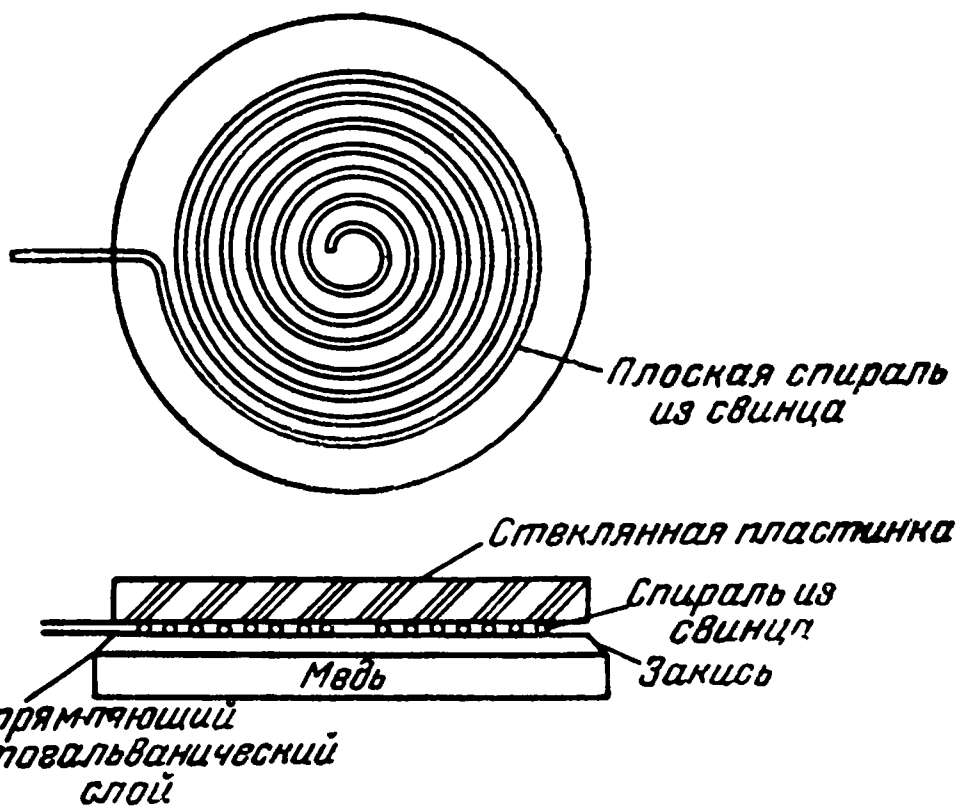


Рис. 8-25. Меднозакисный фотогальванический элемент тылового типа старой конструкции.

в виде редкой плоской спирали из свинцовой проволоки, прижатой стеклянной пластиной (рис. 8-25). Свет, падающий на этот прибор, попадая в промежутки спирального электрода и проходя сквозь слой закиси, генерирован фото-э. д. с. на выпрямляющем контакте. Впоследствии фотоэлементы стали изготовлять с внешним контактом в виде металлической решетки, нанесенной испарением в вакууме с применением трафарета. Подобный же фотоэлемент другой конструкции [Л. 20]

имел внешний контакт в виде полупрозрачного слоя напыленного золота или серебра.

В фотоэлементах, подобных описанным выше, свет должен пройти сквозь оксидный слой, чтобы попасть на фотогальванический запирающий слой. Соответственно этому такие фотоэлементы относят к типу тыловых. В начале 30-х годов были разработаны меднозакисные фотогальванические элементы, в которых выпрямляющий полупрозрачный контакт образуется на *внешней* поверхности слоя закиси меди, так что фото-э. д. с. элемента создается на внешнем контакте. Такие приборы относят к фронтальному типу. Внешний контакт может быть изготовлен осаждением на поверхность закиси меди тонкой пленки золота или платины или путем химического восстановления тонкого поверхностного слоя закиси меди в металлическую медь. Фотоэлементы фронтального типа имеют два важных преимущества перед приборами тылового типа. Во-первых, в последних много света теряется за счет поглощения в слое закиси меди, не доходя до выпрямительного перехода, находящегося за слоем. В фотоэлементах фронтального типа свет не теряется, и потому в них при освещении белым светом развивается ток короткого замыкания примерно на один порядок больше, чем в фотоэлементе тылового типа. Во-вторых, поскольку слой закиси меди поглощает преимущественно синие и зеленые лучи, фотоэлементы тылового типа чувствительны только к красному концу видимого спектра, тогда как фотоэлементы фронтального типа чувствительны ко всем видимым цветам, особенно к синему и зеленому цветам.

8-2-3. Кремниевые фотоэлементы для солнечных батарей

Новейшим и наиболее многообещающим членом семейства фотогальванических элементов является кремниевый фотоэлемент с диффузионным переходом, известный под названием «солнечной батареи» [Л. 21 и 22]. Замечательная особенность этого прибора — преобразование в электроэнергию примерно 7—8% падающего на него солнечного света [1]. Эта величина примерно на один порядок выше, чем у любого из известных ранее типов преобразователей света в электрическую энергию, и да-

[1] К 1957 г. наилучшие экспериментальные кремниевые солнечные батареи имели к. п. д. 11%.

же намного превышает коэффициент использования света при фотосинтезе в листьях растений.

Элемент солнечной батареи можно изготовить, воздействуя на пластину кремния *n*-типа парами треххлористого бора при высокой температуре. В результате проникновения атомов бора в поверхностные слои кремния вокруг тела пластины *n*-типа образуется тонкая пленка кремния *p*-типа. Затем присоединяют электрические контакты к этой оболочке и к внутренней части пластины на участке, где оболочка сошлифована; в результате этого получается плоскостной прибор с *p-n* переходом, расположенным непосредственно под поверхностью. Эффективность этого перехода высока, так как

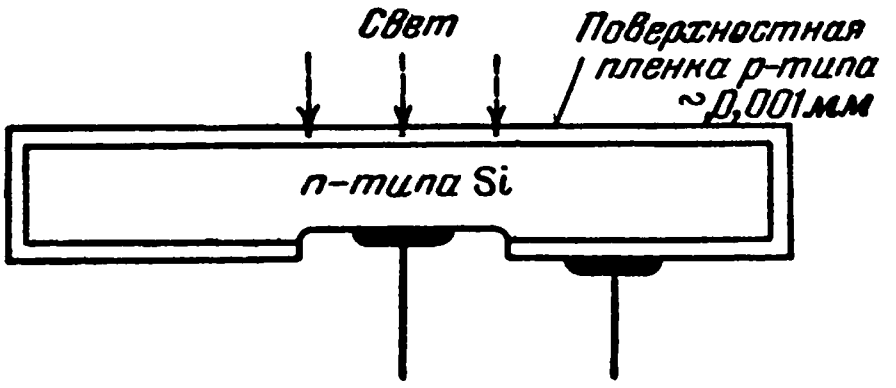


Рис 8-26 Фотогальванический элемент солнечной батареи, в котором *p-n* переход расположен в непосредственной близости к поверхности кремния.

он расположен близко к поверхности и неосновные носители, освобождаемые у поверхности, легко диффундируют к переходу без существенной потери за счет рекомбинации. На рис. 8-26 показан схематический разрез элемента солнечной батареи.

Элемент солнечной батареи дает напряжение холостого хода около 0,5 *в* при освещении прямым солнечным светом в полдень. В цепи, рассчитанной на получение максимальной выходной мощности, напряжение его понижается до 0,3 *в*. В режиме короткого замыкания элемент вырабатывает под действием прямого солнечного света ток примерно от 25 до 30 $ма/см^2$.

Одной из трудностей в производстве солнечных фотоэлементов является создание поверхностного слоя оптимальной толщины. Если переход находится слишком глубоко под поверхностью, то значительная часть неосновных носителей, образовавшихся при поглощении света, успеет рекомбинировать прежде, чем эти носите-

ли дойдут от поверхностных слоев, где они образовались (где произошло поглощение света), до перехода, собирающего носители заряда. С другой стороны, если поверхностный слой слишком тонок, то его продольное сопротивление затрудняет сбор зарядов. Наивыгоднейшая толщина поверхностного слоя — около 10^{-3} *мм*.

Предпочтение кремния в качестве полупроводникового материала для солнечных элементов перед германием обусловлено тем, что *p-n* переход в кремнии позволяет получить при солнечном свете большее напряжение холостого хода и развить более высокую мощность при работе на практическую нагрузку, чем германиевый элемент с такой же площадью. Однако характеристика спектральной чувствительности кремния не простирается так далеко в инфракрасную область, как у германия. Это ограничение несущественно, если батарея освещается солнцем, поскольку максимум интенсивности солнечной радиации приходится на зеленый участок спектра, к которому и кремний и германий приблизительно одинаково чувствительны. Однако при освещении этих фотоэлементов лампой накаливания, у которой максимум излучения лежит в инфракрасной области, германиевый элемент может дать больший к. п. д., несмотря на меньшее выходное напряжение.

8-3. ПРОЧИЕ ФОТОЭЛЕМЕНТЫ РАЗЛИЧНЫХ ТИПОВ

Интересным полупроводником является сульфид таллия, оптические свойства которого явились предметом ряда исследований, описанных в литературе. Из сульфида таллия изготовляются как фотосопротивления [Л. 23], так и фотогальванические элементы [Л. 24]. Для создания фотогальванических элементов слой сульфида таллия образуется на поверхности таллиевого или покрытого таллием основания путем обработки его парами горящей серы. Затем на слой сульфида наносится полупрозрачное покрытие из золота или другого материала, подходящего для создания электрода. Серноталлиевые фотосопротивления изготовляются путем осаждения испарением тонкого слоя сульфида таллия на аквадаговую электродную решетку, нанесенную на внутреннюю поверхность стеклянного баллона. Затем слой сульфида частично окисляется, баллон откачивается и отпаивается. Чувствительность такого элемента,

так же как и серно-свинцового, по-видимому, критически зависит от степени и характера частичного перехода сульфидного слоя в окись. Оба типа серно-таллиевых элементов имеют область спектральной чувствительности, простирающуюся в диапазоне длин волн от 0,5 до 1,5 *мк* с максимумом около 1 *мк*. Измеренный квантовый выход фотосопротивления равен 25.

В современной технической литературе опубликовано много статей, описывающих изготовление и характеристики фотоэлементов из различных оксидов, сульфидов, селенидов, теллуридов и химически простых полупроводниковых веществ с квантовым выходом, много большим единицы. Хотя большинство таких приборов находится еще в стадии лабораторных исследований, все же можно ожидать, что появятся новые приборы, имеющие необычайно хорошие один или несколько параметров, такие, как высокий квантовый выход, очень широкий или очень узкий диапазон спектральной чувствительности, малую инерционность или высокую стабильность в отношении температуры или времени.

8-4. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Как фотосопротивления, так и фотогальванические элементы могут быть выполнены с использованием *p-n* переходов. Однако оба типа фотоэлектрических явлений свойственны в общем случае любому *p-n* переходу и, следовательно, также и выпрямительным контактам между металлом и полупроводником. Взаимосвязь между фотопроводимостью и фотогальваническими свойствами *p-n* переходов станет более ясной, если внимательно изучить поведение семейства электрических и световых характеристик при низких напряжениях. Такое изучение можно выполнить, если построить характеристики в функции, например, кубического корня из напряжения. Такой масштаб дает возможность показать всю характеристику, но в значительно растянутом масштабе в области ниже 1 *в*. На рис. 8-27 показано семейство характеристик фотоэлемента с *p-n* переходом, приведенных на рис. 8-16, построенное в новых координатах, с тщательно воспроизведенными особенностями в области от —1 *в* до +1 *в*.

Первое, что мы замечаем на этом графике, это то, что только кривая темнового тока проходит через на-

чало координат; все остальные кривые пересекают ось напряжений при конечных величинах прямого напряжения. Это семейство кривых показывает, что один и тот же переход по своим характеристикам может служить как фотогальваническим элементом, так и фото-

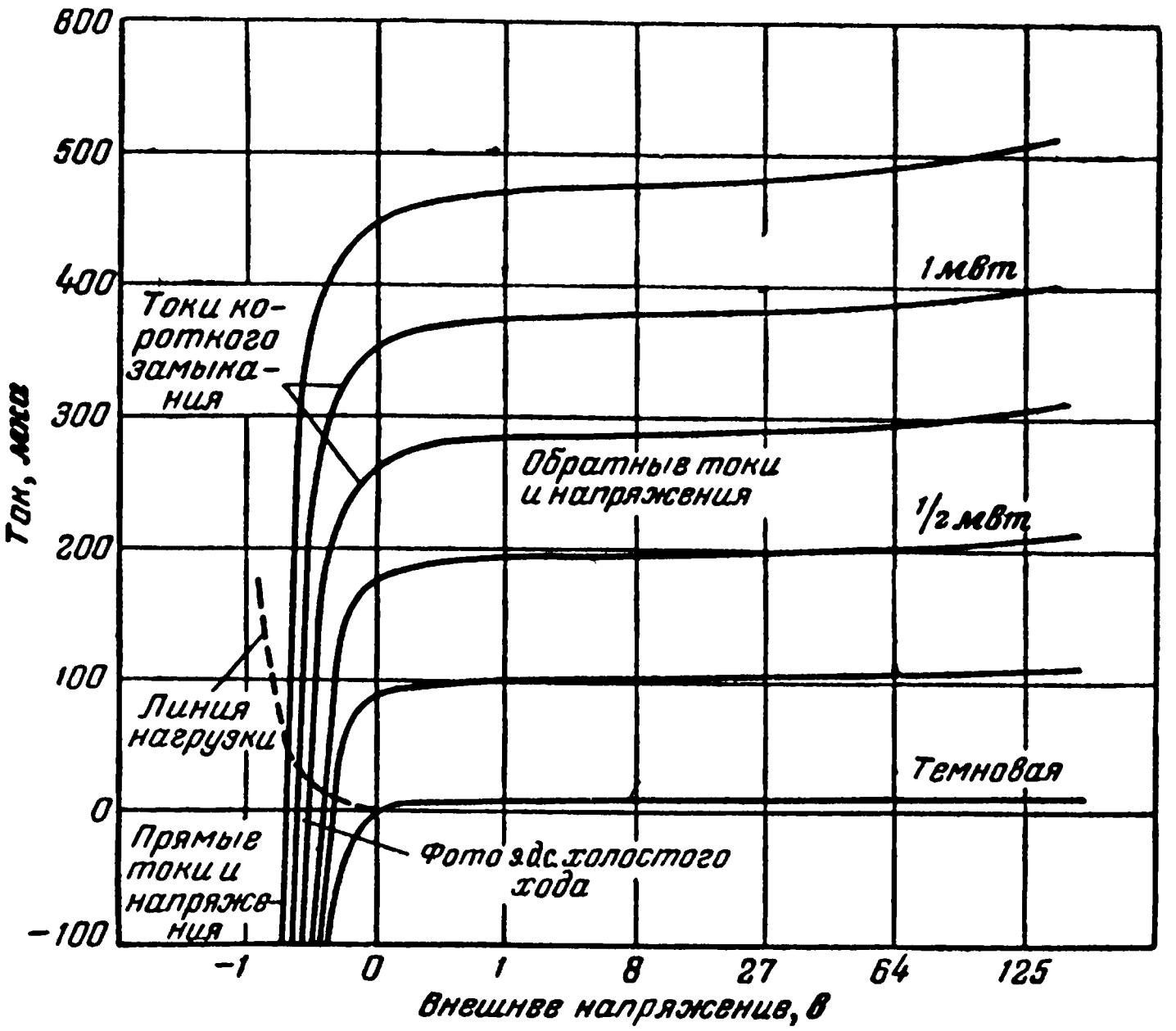


Рис. 8-27 Характеристики фотоэлемента с *p-n* переходом, на абсциссе которых отложены корни кубические из напряжения питания, чтобы выявить важные особенности характеристик при малых напряжениях (от нуля до —1 *в*).

сопротивлением. Если прибор используется как фотогальванический элемент без какого-либо внешнего источника напряжения, то нагрузочная линия проходит через начало координат[1]. Если цепь прибора разомкнуть, то наклон нагрузочной линии будет равен нулю и выходные напряжения холостого хода получатся в точках пересечения кривых с осью напряжений. Если выход прибора замкнут накоротко, то наклон нагрузоч-

[1] Не следует удивляться странной форме нагрузочной линии на рис. 8-27; поскольку масштаб оси напряжений не линейный, и нагрузочная линия получается не прямой, а искривленной, как показано на рисунке.

ной линии равен бесконечности и токи короткого замыкания определятся пересечениями кривых с осью токов. Практически эти токи имеют почти такие же величины, как и при работе прибора в качестве фотосопротивления с приложенным к нему обратным напряжением. Выходные напряжения и токи при конечных сопротивлениях нагрузки определяются точками пересечения кривых с нагрузочной линией, расположенной во втором квадранте. При любой данной величине светового потока нагрузочная линия, соответствующая оптимальному выходу мощности, будет располагаться так, что точка ее пересечения с соответствующей этому световому потоку характеристикой окажется на вершине прямоугольника с максимальной площадью, две смежные стороны которого лежат на осях тока и напряжения. При этом фотогальванический элемент находится в довольно аномальном режиме генерирования обратного тока при прямом напряжении. Этот режим неустойчив и может поддерживаться только за счет расходования световой энергии внутри фотоэлемента.

Если элемент используется как фотосопротивление, то нагрузочную линию надо сместить от начала координат настолько, чтобы пересечь ось напряжений в точке, соответствующей надлежащей величине приложенного обратного напряжения. Анализ характеристик прибора и его цепи выполняется затем обычным образом.

Сравнивать и истолковывать характеристики фотоэлементов различных видов и типов трудно, хотя бы потому, что различные фотоэлементы чувствительны к разным частям спектра. Поэтому фотоэлементы, имеющие сравнимую чувствительность для видимого света, обнаруживают громадную разницу в характеристиках при работе с лампами накаливания, у которых максимум излучения приходится на инфракрасный участок. В настоящей главе был введен в качестве характеризующего параметра квантовый выход; однако следует иметь в виду, что при равной энергетической интенсивности излучения синего и красного цветов в потоке последнего проходит вдвое больше фотонов через единицу площади, нормальной к этому потоку.

Сравнение характеристик осложняется также необходимостью учета оптимального распределения входно-

Сравнение свойств

Тип	Вид	Активная поверхность, $см^2$	Длина волны в максимуме чувствительности, *мк*
Электровакуумный фотоэлемент	Электронный	~1	0,8* **
	Ионный		0,8
	Фотоэлектронный умножитель		Та же, что и для обычного электронного фотоэлемента с тем же фотокаскадом
Фотосопротивление	Селеновый [Л.1 и 3]	От 0,1 до ~6	—
	Серно-свинцовый [Л.4]		1—2,6 в зависимости от термообработки
	Точечный германиевый [Л. 7,8]	$\sim 10^{-3}$	1,5
	Германиевый с *p-n* переходом [Л.13]	~0,005	1,5
	Серно-таллиевый [Л. 2]	0,1—~6	0,9
Фотогальванический элемент	Селеновый [Л.1]	6	—
	Меднозакисный фронтальный	0,6	0,5
	Меднозакисный тыловой [Л.18 и 19]	6	0,65

Таблица 8-1

фотоэлементов

Пороговая длина волны, *мк*	Частота, при которой чувствительность спадает на 3 *дб* от чувствительности при низкой частоте, *кгц*	Темновое сопротивление переменному току в нормальном рабочем режиме, *Мом*	Чувствительность, технич единиц
1,2	—	Примерно бесконечно	20 *мка/лм* при нагрузке 1 *Мом*
1,2	15		90 *мка/лм* при нагрузке 1 *Мом*
	—		В зависимости от числа каскадов
1	5	0,5—10	1 *ма/лм*
3	5	0,2—5	—
1,8	Более 200 для светового пятна, падающего в пределах 0,25 *мм* от центра	—	100 *ма/лм* при нагрузке 20 *ком* и цветовой температуре 2 870° С
1,8	200 для светового пятна падающего в пределах 0,125 *мм* от центра	10	30 *ма/лм* при цветовой температуре 2 870° С
1,3	—	0,1—100	—
1	0,2—6	Изменяется в зависимости от интенсивности падающего светового потока	80 *мка/лм* в режиме короткого замыкания
0,65	10—15		500 *мка/лм* в режиме короткого замыкания
—	Более 6		75 *мка/лм* в режиме короткого замыкания

Тип	Вид	Активная поверхность, $см^2$	Длина волны в максимуме чувствительности, *мк*
Фотогальванический элемент	Германиевый с *p-n* переходом	0,005	1,5
	Кремниевая солнечная батарея [Л.20 и 21]	1—~20	1

* Спектральная характеристика типа S-1.

** Известно примерно полдюжины форм спектральных характеристик, указы Тип S-1 имеет максимальную чувствительность для красного цвета.

го светового потока по площади фотоэлемента. У некоторых фотоэлементов выход больше при равномерном распределении света по активной поверхности; другие, как, например, точечные и плоскостные фототранзисторы, лучше используют поток света, если он сфокусирован в пятно малого диаметра и направлен на наиболее чувствительную часть поверхности элемента. Точно так же инерция и предельная верхняя частота фотоэлемента с переходом зависят от того, на каком расстоянии от перехода падает свет.

Некоторые сравнительные характеристики приведены в табл. 8-1. Для полноты указаны данные некоторых электровакуумных фотоэлементов. Параметры некоторых фотоэлементов, приводимые в литературе, иногда значительно различаются; в таких случаях при составлении таблицы приходилось усреднять опубликованные данные, чтобы получить цифры, представляющие типичные характеристики описываемых образцов. Приведенные в таблице ссылки на литературу соответствуют ссылкам в тексте настоящей главы.

ЗАДАЧИ

1. Фотосопротивление, имеющее характеристики, приведенные на рис. 8-1, равномерно освещаются по всей верхней поверхности светом, который экспоненциально поглощается в толще полупроводника

Продолж. табл. 8-1

Пороговая длина волны, *мк*	Частота, при которой чувствительность спадает на 3 *дб* от чувствительности при низкой частоте, *кгц*	Темновое сопротивление переменному току в нормальном рабочем режиме, *Мом*	Чувствительность, технич единиц
1,8	~200 для светового потока, падающего в пределах 0,125 *мм* от области перехода	Изменяется в зависимости от интенсивности падающего светового потока	400 *мка/мвт* в режиме короткого замыкания
1,1	—		30 *ма/см*^{2} в режиме короткого замыкания на солнечном свете

ваемых изготовителями фотоэлементов и обозначаемых как типы S-1, S-2 и т. д

и интенсивность которого такова, что удельное сопротивление поверхностного слоя полупроводника понижается до одной десятой его темновой величины. Интенсивность света на глубине 1 *мм* от поверхности равна $1/e$ его величины на поверхности. Определите и постройте график электропроводности полупроводника между электродами как функции толщины полупроводника в пределах от 0 до 5 *мм* для затемненного и освещенного состояний.

2. Десять элементов кремниевой солнечной батареи соединены последовательно и при освещении питают последовательно включенную нагрузку. Предположим, что по какой-либо причине свет падает только на девять элементов, оставляя десятый в темноте. Как изменится ток в цепи?

ЛИТЕРАТУРА

1. Barnard G. P., The Selenium Cell, R. R. Smith, Inc., New York, 1930.
2. Zworykin V. K., Ramberg E. G., Photoelectricity, Wiley, New York, 1949.
3. German patent No. 219574, February 18, 1909.
4. Pakswer S., Lead Sulfide Photoconductive Cells, Electronics, May 1949, vol. 22, p. 111.
5. Sosnowski L., Lead Sulfide Photoconductive Cells, Nature, 1947, vol. 159, p. 818.
6. Cashman R. J, Development of Sensitive Lead Sulfide Photoconductive Cells for Detection of Intermediate Infra-red Radiation, O.S.R.D. Report 5998, October 31, 1945.
7. Shive J. N., A new Germanium Photoresistance Cell, Physical Review, 1949, vol. 76, p 575.
8. Shive J. N., The Phototransistor, Bell Laboratories Record, 1950, vol. 28, p. 337.

9. Fritts E. E., On a New Form of Selenium Photocell, Proceedings of the American Association for the Advancement of Science, 1883, vol. 33, p. 97; American Journal of Science, 1883, vol. 26, p. 465.

10. Kingsbury E. F., Ohl R. S., Photoelectric Properties in Ionically Bombarded Silicon, Bell System Technical Journal 1952, vol. 31, p. 802

11. Benzer S., Excess-Defect Germanium Contacts, Physical Review, 1947, vol. 72, p. 1267.

12. Pietenpol W. J., p—n Junction Rectifier and Photocell, Physical Review, 1951, vol. 82, p. 120.

13. Shive J. N., Properties of M-1740 p—n Junction Photocells, Proceedings of the Institute of Radio Engineers, 1952, vol. 40, p. 1410.

14. Goucher F. S., Measurement of Hole Diffusion in n-Type Germanium, Physical Review, 1951, vol. 81, p. 475.

15. Shive J. N., Properties of Germanium Phototransistors, Journal of the Optical Society of America, 1953, vol. 39, p. 243.

16. Adams W. G., Day R. E., The Action of Light on Selenium, Proceedings of the Royal Society, 1877, vol. A25, p. 113.

17. Fritts C. E. (см. [Л. 9] этой главы).

18. Bergmann L., Über eine neue Selen-Sperrschicht Photozelle, Physikalische Zeitschrift, 1931, vol. 32, p. 286.

19. Grondahl L. O., The Copper-Cuprous-Oxide Rectifier and Photoelectric Cell, Reviews of Modern Physics, 1933, vol. 5, p. 141.

20. Lange B, Die Photoelemente und ihre Anwendung, Johann Barth, Leipzig, 1936.

21. Chapin D. M., Fuller C. L., Pearson G. L., A New Silicon p—n Junction Photocell for Converting Solar Radiation into Electrical Power, Journal of Applied Physics, 1954, vol. 25, p. 676.

22. Prince M. B., Silicon Solar Energy Converters, Journal of Applied Physics, 1955, vol. 26, p. 534.

23. Hippel A., Thallous Sulfide Photoconductive Cells, Journal of Chemical Physics, 1946, vol. 14, p. 355.

24. Nix F. C., Treptow A. W., A Thallous Sulfide Photo EMF Cell, Journal of the Opical Society of America, 1939, vol. 29, p. 457.

25. Рашба Э. И., Первое Всесоюзное совещание по фотоэлектрическим и оптическим явлениям в полупроводниках, ЖТФ, 1958, 12, 2696—2706.

26. Коломиец Б. Т., Фотосопротивления, Труды первой межвузовской конференции по современной технике диэлектриков и полупроводников, июнь 1956, Изд Ленинградского электротехнического института имени В. И. Ульянова (Ленина), 1957, 246—255.

27. Коломиец Б. Т и Пратусевич С. Г., Фотосопротивления из селенида кадмия, «Радиотехника и электроника», 1956, 1 8, 1174—1176.

ГЛАВА ДЕВЯТАЯ

ПРИМЕНЕНИЕ ПОЛУПРОВОДНИКОВЫХ ФОТОЭЛЕМЕНТОВ

Хотя открытие внутреннего и фотогальванического эффектов в полупроводниках произошло за несколько десятилетий до открытия фотоэлектронной эмиссии и появления электровакуумных фотоэлементов, именно последние применяются почти во всех тех случаях, когда требуются фотоэлектрические приборы. Только в устройствах, работа которых основана на использовании фотогальванического эффекта, полупроводниковые фотоэлементы сохранили исключительную монополию.

Причины предпочтения во многих применениях электровакуумных фотоэлементов полупроводниковыми различны и в некоторых случаях неустранимы. До появления фотоэлементов с *p-n* переходом электровакуумные фотоэлементы были единственными приборами, обеспечивавшими в широких пределах строгую пропорциональность выходного тока световому потоку, необходимую, например, при воспроизведении звука в кино. В других случаях электровакуумные фотоэлементы предпочитают из-за частотных ограничений, свойственных полупроводниковым приборам. Применение многих полупроводниковых приборов, кроме того, связано с проблемами старения, быстрой усталости и чрезвычайно сильных температурных изменений темнового тока и чувствительности.

Несмотря на эти недостатки, существует много устройств, в которых с успехом используются полупроводниковые фотоэлементы. Некоторые из них описаны ниже.

9-1. ЦЕПИ ПОСТОЯННОГО ТОКА

На постоянном токе фотоэлементы используются просто для того, чтобы по величине тока определить интенсивность постоянного светового потока. Для этой цели пригодны фотосопротивления, но они требуют источников питания. Поэтому в большинстве таких устройств применяются фотогальванические элементы, так как они сами являются источниками э. д. с. Существуют различные переносные приборы для измерения света, состоящие из фотоэлемента и измерительного

прибора постоянного тока. Так, к этой категории приборов относятся люменометры, люксметры и экспонометры, используемые в фотографии.

В люксметрах обычно желательно, чтобы шкала отсчета индикаторного прибора была линейной по отношению к входному световому потоку. Для этого внутреннюю цепь прибора устраивают так, чтобы сопротивление нагрузки фотоэлемента было мало, обеспечивая этим линейную зависимость тока от падающего светового потока в сравнительно широком диапазоне интенсивностей света (рис. 8-22). С другой стороны, для фотоэкспонометров часто желательна нелинейная характеристика, так как этими приборами приходится измерять интенсивности света, различающиеся в сотни раз.

9-2. ИСПОЛЬЗОВАНИЕ ФОТОЭЛЕМЕНТОВ В КАЧЕСТВЕ УПРАВЛЯЕМЫХ СВЕТОМ КЛЮЧЕЙ ПРИ ВЫСОКИХ УРОВНЯХ СИГНАЛОВ

Фотоэлементы, используемые в качестве ключей, управляемых светом при высоких уровнях сигналов, просто различают присутствие и отсутствие света. Когда фотоэлемент освещен, его выходной ток используется для осуществления некоторого действия, которое не может быть выполнено, если фотоэлемент не освещен. Так, например, в блокирующих устройствах для защиты оператора, управляющего опасными механизмами, источник света и фотоэлемент должны быть расположены так, чтобы, пока руки рабочего находятся в опасной зоне, световой поток на фотоэлемент не падал и цепь, управляющая работой механизма, была разомкнута. В фотоэлектрических счетчиках источник света и фотоэлемент располагают по обеим сторонам ленты счетного конвейера: предметы, движущиеся по конвейеру, на мгновение прерывают световой луч, и счетчик срабатывает. Различные варианты этого метода включения и выключения используются для счета машин при изучении уличного движения, для открывания дверей при приближении людей, включения сигналов тревоги при проникновении воров или при пожарах, при появлении дыма, а также для включения уличных фонарей и световых реклам при наступлении темноты. Эти применения общеизвестны. Литература [Л. 1 и 2] по подобному использованию фотоэлементов весьма обширна.

На рис. 9-1 и 9-2 показаны в упрощенном виде примеры использования фотоэлементов для считывания информации с перфоленты стартстопного телеграфного аппарата. В считывающем устройстве лента протягивается между источником света и группой фототранзисторов, установленных так, что каждый фотоэлемент считывает с ленты лишь один разряд. Такое считывание

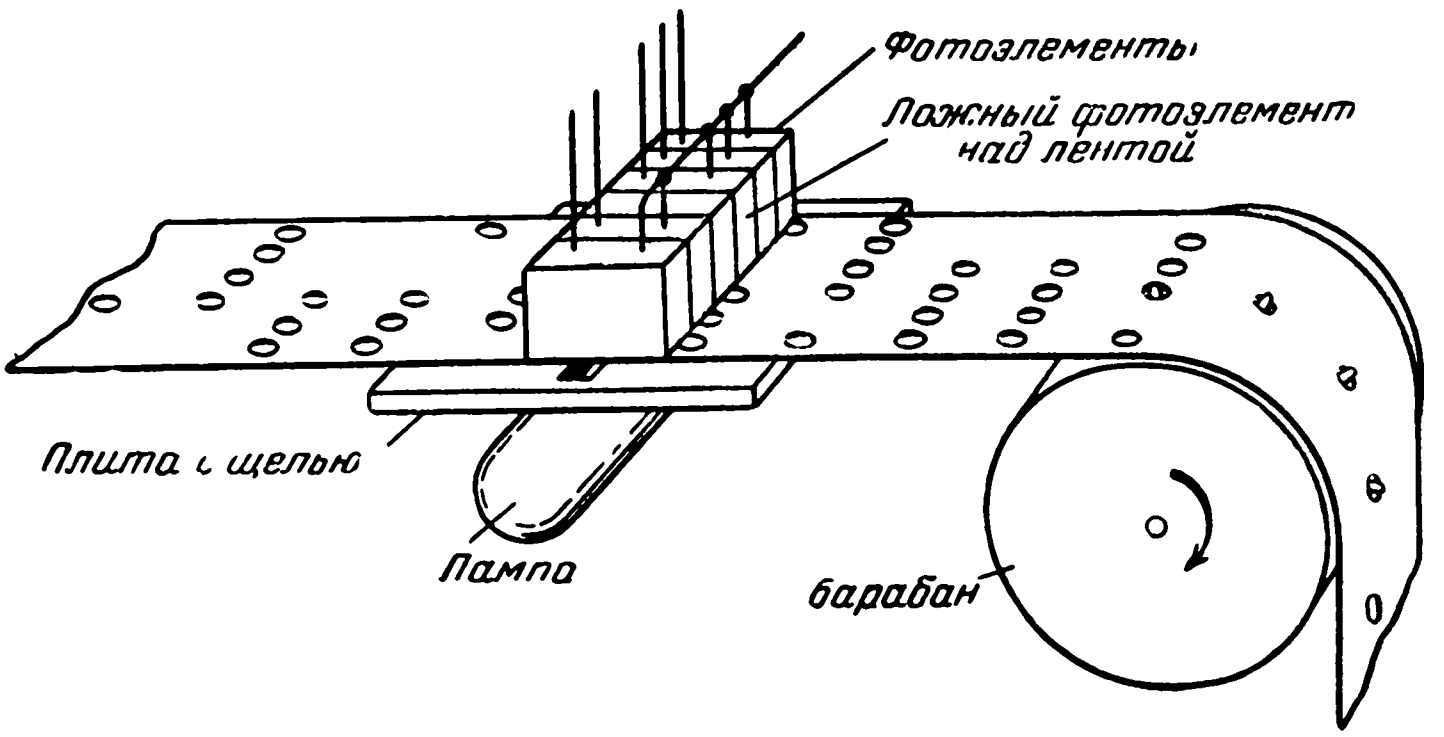


Рис. 9-1. Схема считывающего устройства Свет, проходящий через щель и отверстия в перфорированной ленте, поступает к фотоэлементам с *p-n* переходами.

может осуществляться быстрее, чем с помощью металлического контакта через дырки на ленте; кроме того, такое устройство полностью свободно от попадания бумажных волокон между металлическими контактами и лента меньше изнашивается.

В кодирующем устройстве [Л. 3] каждая клавиша, соответствующая букве или символу, соединена с планкой, имеющей кодовые выступы, как показано на рис. 9-2. При нажиме на клавишу кодовые выступы пре-

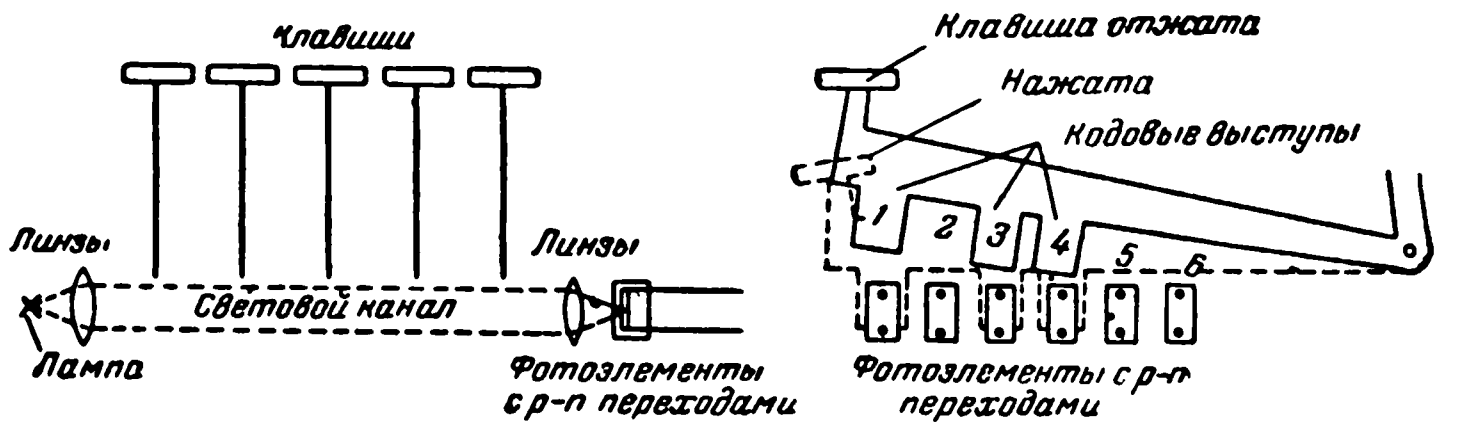


Рис. 9-2. Схема кодирующего устройства с фотоэлементами. При перекрытии световых каналов образуются кодовые сочетания, соответствующие отдельным буквам или символам.

рывают пучки света, проходящие под клавишными планками от одного конца клавиатуры к другому. Расположение кодовых выступов на каждой клавише различно, и ее нажатие дает разные кодовые комбинации открытых и закрытых световых каналов. В результате при нажиме на клавишу механическое перфораторное устройство, управляемое выходными сигналами фотоэлемента, пробивает кодовые отверстия в ленте, которая затем вводится в трансмиттер телеграфного аппарата для передачи обычным способом. В этом приборе первоначально использовались электровакуумные фотоэлементы, но они слишком велики для установки их в один ряд вдоль выходной боковой стороны клавиатуры; поэтому при их размещении требовались некоторые оптические приспособления. На рис. 9-2 показано возможное выполнение считывающего устройства, в котором в один ряд расположены фотоэлементы с p-n переходом. Наибольшее число различных символов, которое может быть закодировано с помощью N каналов, равно 2^N. Поэтому вполне достаточно по кодовой емкости пяти каналов, чтобы однозначно закодировать все буквы английского алфавита. А если добавить шестой канал, то можно закодировать также все цифры и несколько часто употребляемых символов и буквенных сочетаний.

Интересное применение фотоэлементы нашли в качестве ключей для кодирования углового положения главного вала при повторении этого углового положения на одном или большем числе вынесенных вторичных валов (в сервоприводах типа «электрического вала») [Л. 4]. На рис. 9-3 показана схема кодирующего устройства такой системы. Главный вал снабжен прозрачным кодовым диском, на одной стороне которого нанесены кодовые обозначения по двоичной системе в виде трафарета, состоящего из непрозрачных секторов. С одной стороны диска расположен источник света, с другой — ряд фототранзисторов с переходами, на которые попадает свет от кодового трафарета через узкую считывающую щель. Если нужно закодировать угловое положение вала, то производится вспышка источника света (например, импульсной газоразрядной лампы). Выходные сигналы фотоэлементов, представляющие собой комбинации единичных импульсов и нулей

двоичной системы, запоминаются каждый в своей ячейке накопительного устройства с параллельными входами и затем считываются последовательно во времени. Единицы и нули (по двоичной системе) передаются затем по линии к сервоустройствам вторичных валов, где

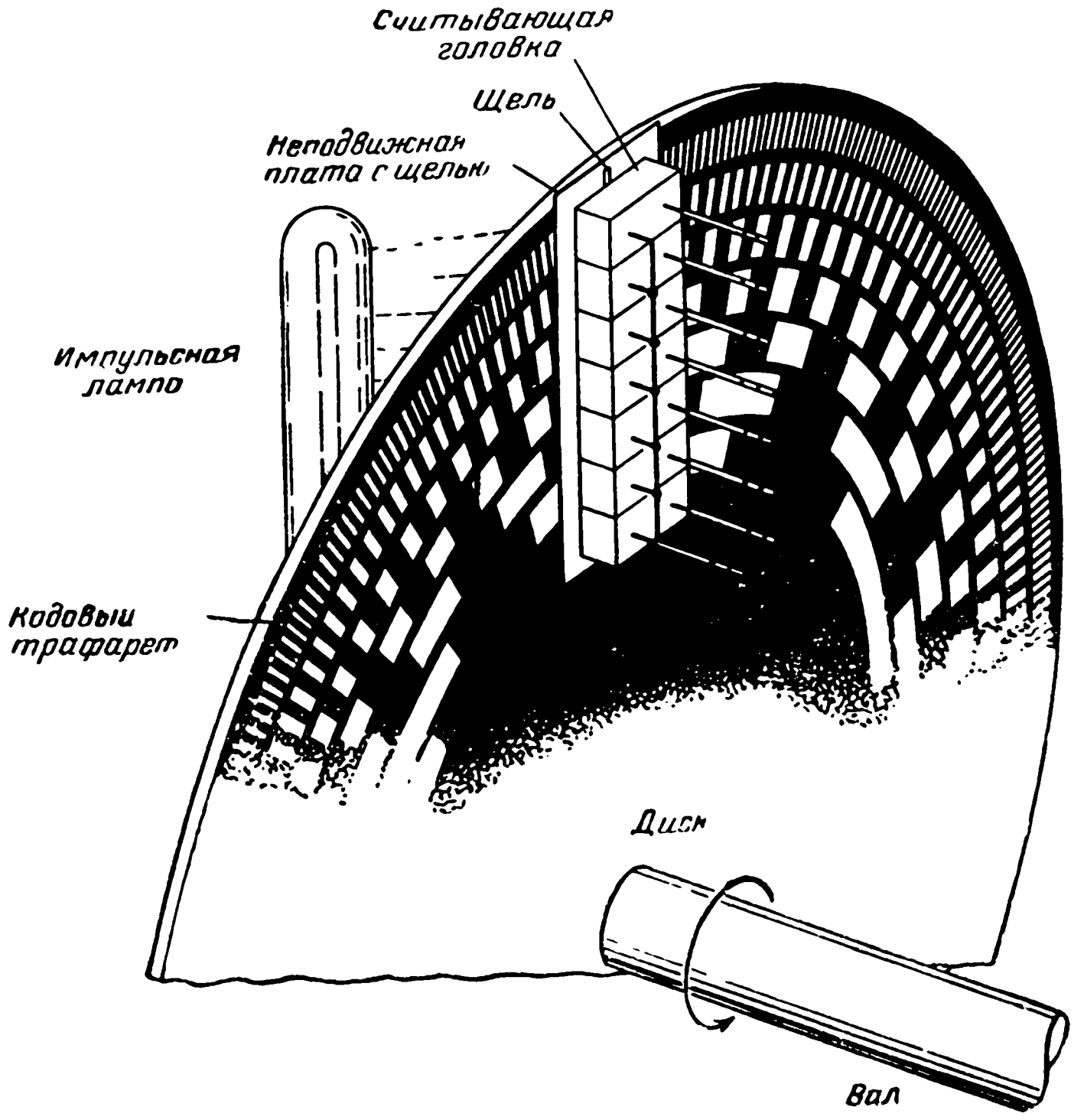


Рис. 9-3 Кодирующее устройство, указывающее угловое положение вала с помощью считывающей головки с фототранзисторами, воспринимающими свет, проходящий через светлые места кодового трафарета, нанесенного на диске

они служат для сравнения фактического положения вторичного вала с требуемым, что выполняется с помощью подобного же кодирующего устройства на вторичном валу. Результирующий сигнал рассогласования управляет следящей системой, которая ставит вторичный вал в требуемое положение.

Нанесенный на кодовый диск двоичный код из восьми разрядов позволяет определить положение вала

с угловой ошибкой менее 1°. Описана возможность применения 13-разрядного кода [Л. 5 и 6], допускающего определение положения диска с ошибкой менее двух угловых минут с помощью кодового диска диаметром около 25 *см*. На таком кодовом диске черно-белые секторы зоны с самыми мелкими делениями образуют кольцо, состоящее из просветов шириной ~0,05 *мм*,

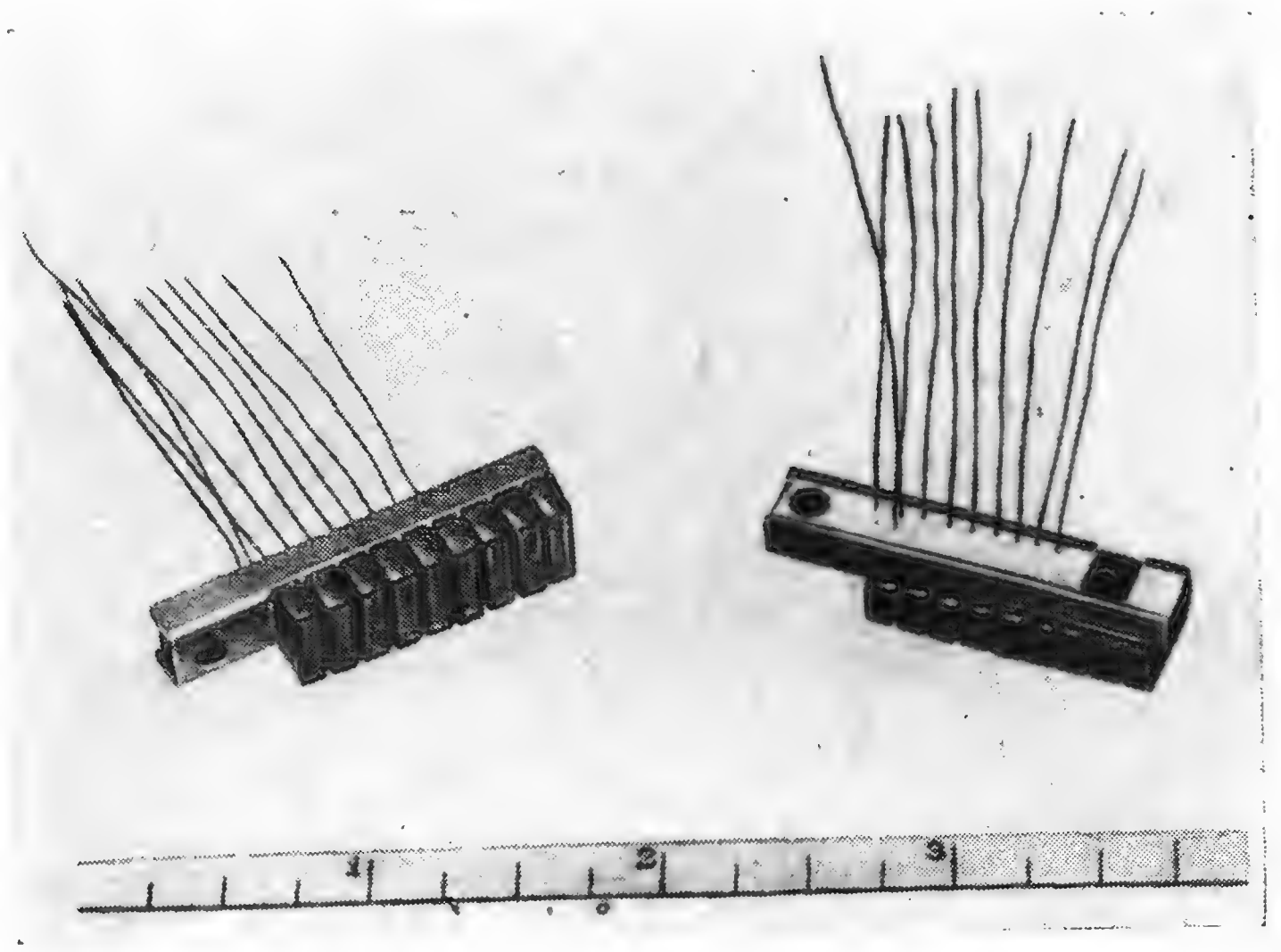


Рис 9-4. Головка из девяти фотоэлементов с *p-n* переходами, считывающая угловое положение диска

разделенных черными штрихами такой же ширины. Код считывается фотоэлементами через неподвижную щель шириной ~0,025 *мм*. На рис. 9-4 показана фотография экспериментального девятиразрядного фотоэлемента считывающего устройства такого назначения.

Интересным примером использования фотоэлементов в режиме ключей при несущей частоте 400 *гц* является фотоэлектрическое карточное адресное устройство, используемое в общегосударственной телефонной системе прямого вызова [Л. 7]. Вызов осуществляется оптико-электромеханическим устройством, применяемым для запоминания и считывания информации, необходимой для направления телефонного вызова по системе

дальней связи, проходящего через промежуточные телефонные станции, на которых установлены подобные вызывные устройства. Вся информация содержится в серии из нескольких сотен прямоугольных металлических карточек, перфорированных согласно коду и установленных на ребро на краю карточного магазина вызывного устройства. На всех карточках имеется опорная сетка из 118 прямоугольных дырочек. Когда карточки находятся в магазине, все 118 световых каналов проходят

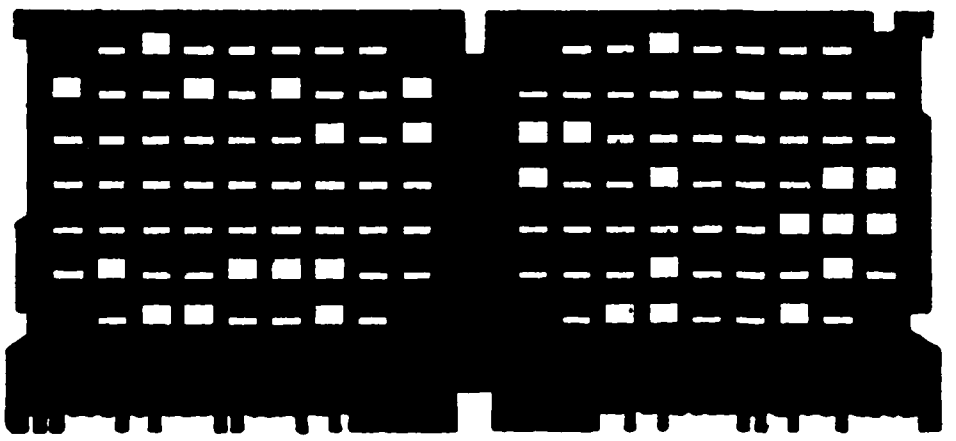


Рис. 9-5 Кодовая карточка, информация на которой нанесена посредством увеличения площади некоторых отверстий

насквозь стопку карточек. На рис. 9-5 показан силуэт одной из кодовых карточек. Информация регистрируется на карточках не этими 118 основными отверстиями, а комбинацией тех отверстий на карточке, размер которых увеличен, как показано на рисунке. После выбора какой-нибудь определенной карточки она смещается примерно на 6 *мм* ниже остальных карточек в пачке, так что все световые каналы перекрываются, кроме тех, которые соответствуют пробитым в данной карточке увеличенным отверстиям. Фотоэлементы, установленные в узлах прямоугольной решетки на выходном конце магазина с карточками, при этом воспримут свет только от тех каналов, которые остались открытыми; в дальнейшем эта информация преобразуется в команды, которые производят соответствующую коммутацию телефонных каналов, ведущих на следующий коммутационный узел по пути к конечному назначению.

Выбор и смещение карточки осуществляются следующим образом Кодовая карточка имеет на нижнем краю множество кодовых выступов, видных на силуэте карточки (рис. 9-5). Комбинации выступов различны для каждой карточки в стопке. Эти выступы опирают-

ся на кодовые стержни, проходящие под всей стопкой и поддерживающие все карточки. Когда нужно выбрать какую-нибудь карточку согласно номеру, набранному абонентом, осуществляющим вызов, кодовые стержни, соответствующие комбинации выступов данной карточки, опускаются, так что только одна выбранная карточ-

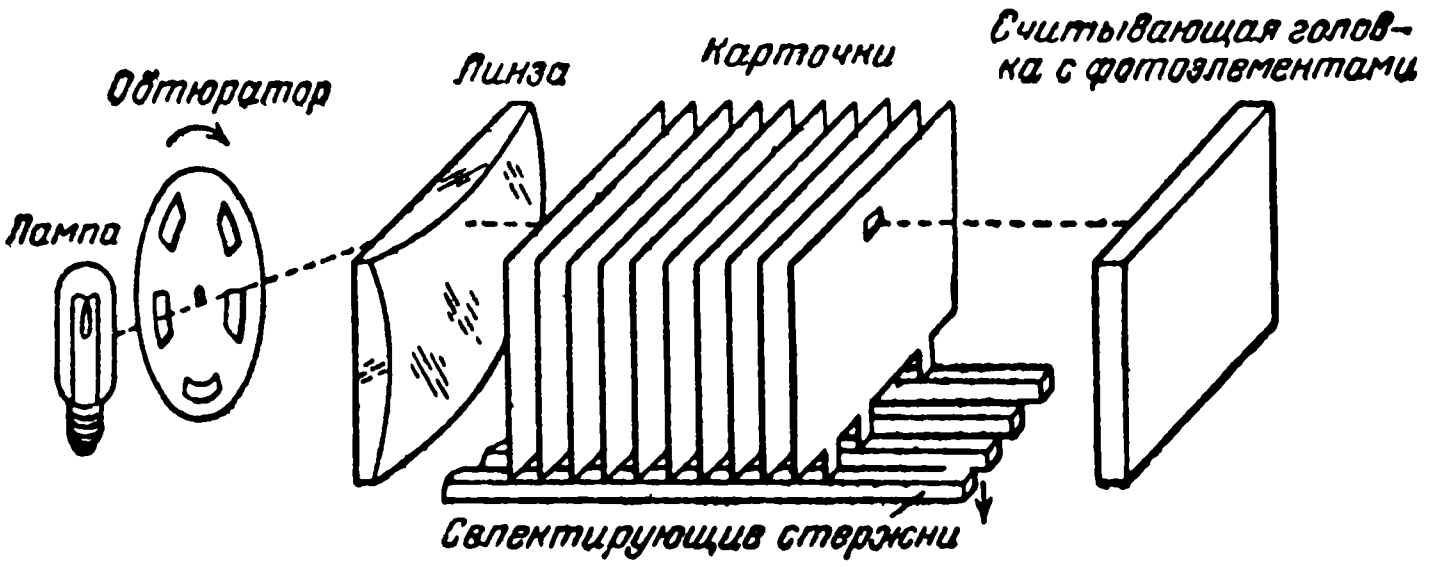


Рис. 9-6. Упрощенная схема фотоэлектрического адресного устройства.

ка смещается вниз, в каком бы месте стопки она ни находилась. Взаимодействие этих частей устройства показано на рис. 9-6.

На рис. 9-7 показана цепь одного из фототранзисторов считывающего устройства. Свет, проходящий через каналы в стопке карточек, прерывается обтюраторным диском с частотой 400 *гц*. Выходной сигнал фотоэлемента, соответствующего освещенному каналу, содержит поэтому составляющую частоты 400 *гц*, которая усиливается однокаскадным транзисторным усилителем. На-

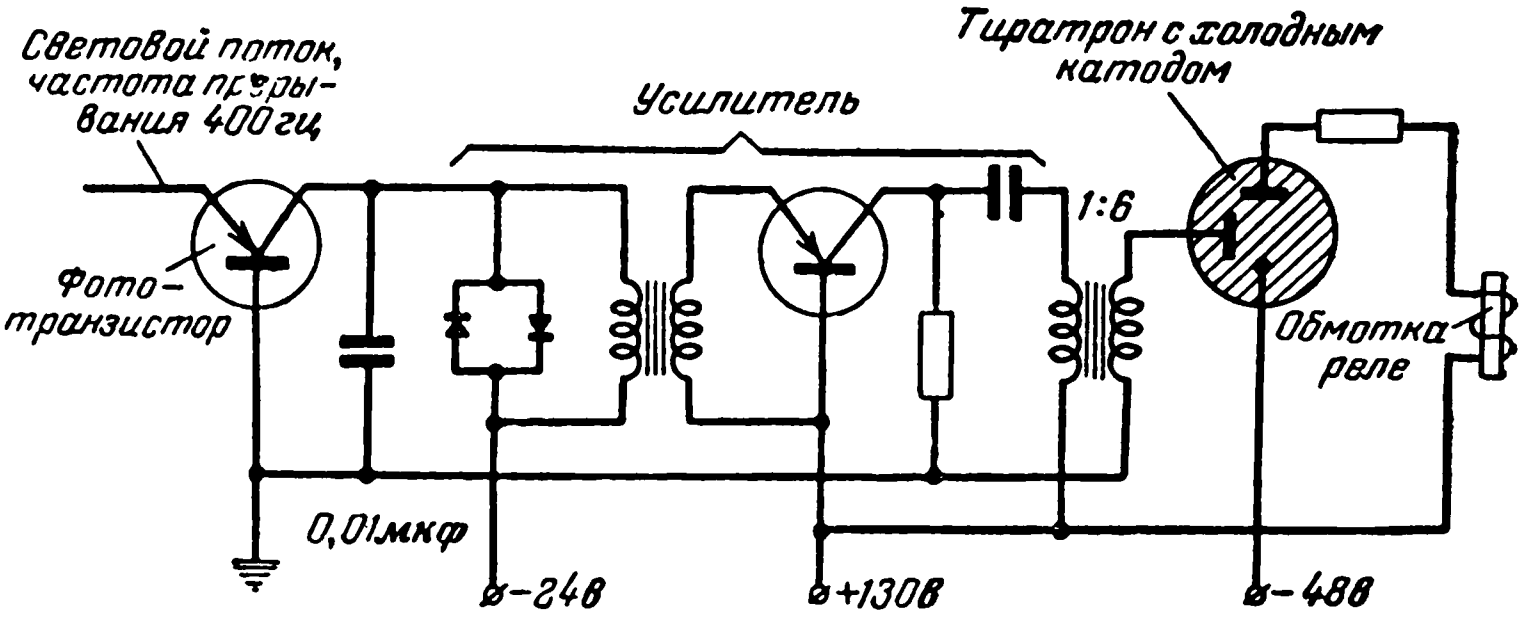


Рис 9-7 Принципиальная схема одного из каналов фотоэлектрического адресного устройства Сигнал на выходе освещенного фотоэлемента служит для зажигания тиратрона, в цепь которого включена обмотка реле.

пряжение с выхода усилителя подается на трансформатор с коэффициентом трансформации 1:6 и далее на управляющий электрод тиратрона с холодным катодом, в цепи которого имеется реле, срабатывающее в этот момент. В цепи фотоэлемента предусмотрен конденсатор емкостью 0,01 *мкф* для шунтирования коммутационных помех, но он мало влияет на сигнал с частотой 400 *гц*. Транзистор защищен от возможных перегрузок симметричным варистором.

9-3. ЦЕПИ ПЕРЕМЕННОГО ТОКА

Фотоэлементы в цепях переменного тока применяются при воспроизведении звукозаписи с кинолент, генерировании музыкальных тонов с помощью звуковых дорожек на тондисках в электромузыкальных синтезаторах и при передаче неподвижных изображений (фототелеграфия). Эти случаи описаны в литературе [Л. 8 и 9], и поэтому мы ограничимся лишь некоторыми замечаниями.

В цепях переменного тока чаще всего используются электровакуумные фотоэлементы. Для хорошего воспроизведения звукозаписи и генерирования музыкальных тонов необходимо выполнение двух требований: линейности выходного электрического сигнала по отношению к входному световому сигналу и постоянства чувствительности в области звуковых частот Электровакуумные фотоэлементы обладают этими свойствами, а полупроводниковые фотоэлементы неполностью удовлетворяют этим требованиям. Но полупроводниковые фотоэлементы чувствительнее электровакуумных на один или несколько порядков, и поэтому они все-таки применяются в ряде случаев в режиме переменного тока.

В гл. 8 говорилось, что, уменьшая нагрузочное сопротивление в цепи фотогальванического элемента, можно расширить линейный участок его характеристики на более широкий интервал интенсивностей входного светового потока. Этот прием используется при работе фотогальванических элементов в оптических звукоснимателях. Однако ограниченная полоса частот постоянной чувствительности этих приборов требует частотной коррекции, осуществляемой схемными методами.

По-видимому, серно-свинцовые полупроводниковые фотоэлементы могут быть изготовлены с характеристи-

ками, в основном соответствующими требованиям для хороших звукоснимателей. С помощью надлежащих технологических приемов был изготовлен [Л. 10] элемент, имеющий спад частотной характеристики на 3 *дб* при ~4 *кгц*. Амплитудная характеристика этого элеменга линейна в достаточно широком диапазоне над уровнем шумов, а область чувствительности в инфракрасном участке спектра пригодна для работы прибора с лампой накаливания в качестве источника света. Чувствительность этого фотоэлемента к инфракрасным лучам допускает работу при сравнительно низкой температуре нити накала в условиях значительной тепловой инерции, чтобы исключить фон удвоенной частоты при питании накала от сети промышленной частоты (в США — 60 *гц*).

Фототранзисторы с *p-n* переходом линейны, высокочувствительны и их частотные характеристики соответствуют требованиям высококачественного воспроизведения звука. Однако ввиду новизны этих приборов возможности работы их в цепях переменного тока еще далеко не изучены.

9-4. ГЕНЕРИРОВАНИЕ ЭНЕРГИИ

Сравнительно недавно были сконструированы фотогальванические приборы, достаточно эффективные для того, чтобы фотоэлектрическое образование солнечного света оказалось экономически выгодным для создания первичного источника малой и средней мощности. Все же, по-видимому, потребуется долгое время, прежде чем прямое преобразование солнечного света сможет в какой-то мере конкурировать с тепловыми или атомными электростанциями. Общее мировое потребление электроэнергии в 1955 г. примерно равнялось мощности солнечного излучения, падающего на площадь 600 $км^2$ на географической широте и в климатических условиях штата Аризона. Однако в наши дни известны случаи, когда использование солнечных батарей в качестве источников энергии экономически оправдано.

Так, например, солнечные батареи могут работать в переносных приемниках, в частности в транзисторных, потребляющих небольшую мощность, а требуемое напряжение не превышает нескольких вольт. Солнечные батареи могут найти применение также в выносных те-

леметрических метеорологических станциях, размещенных в недоступных местах, и в промежуточных усилителях телефонных и телеграфных линий, подвод питания к которым обошелся бы слишком дорого. В тех случаях, когда необходимо непрерывное питание, может быть придан аккумулятор, который заряжается в солнечное

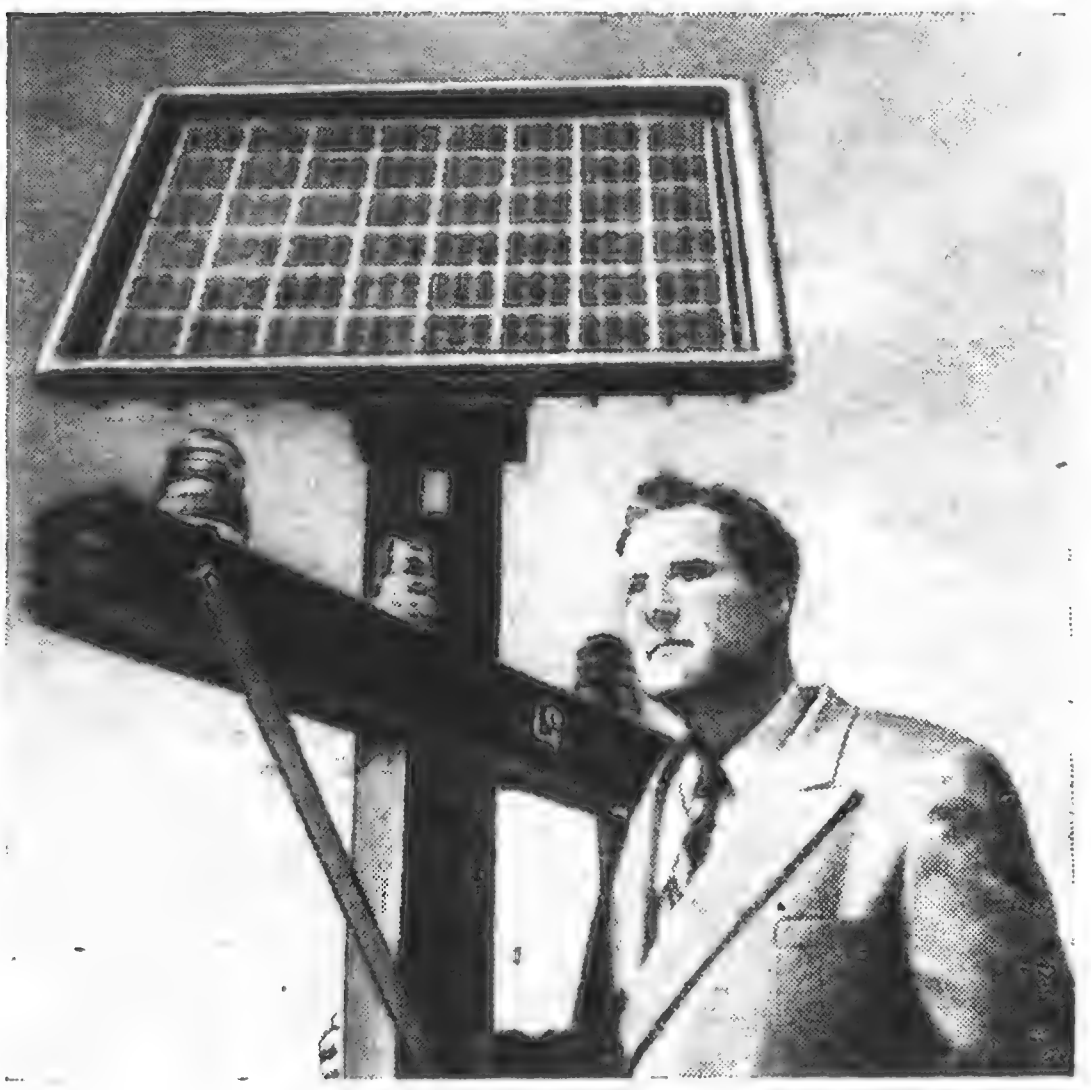


Рис. 9-8. Солнечная батарея из 432 кремниевых элементов, дающая на прямом солнечном свете около 500 *ма* при 22,5 *в*

время суток и обладает достаточной емкостью для питания станции ночью или в облачные дни.

Производились опыты применения солнечной батареи для питания промежуточного усилителя на транзисторах, новой многоканальной телефонной системы. Батарея состояла из 400 солнечных фотоэлементов, площадью около 4,5 $см^2$ каждый, элементы соединены были в последовательные и параллельные цепи; при прямом солнечном свете батарея давала 0,5 *а* при 22,5 *в*. На рис. 9-8 показана эта батарея наверху телеграфного столба, на котором установлен промежуточный усилитель. Батарея наклонена в южную сторону под углом

к горизонту, соответствующим географической широте местности. Электроэнергия, вырабатываемая этой батареей, превышает расход энергии, потребляемой усилителем; в солнечные дни этот избыток используется для зарядки аккумулятора, установленного на том же столбе. Ночью и в пасмурные дни усилитель питается от аккумулятора.

Солнечные батареи — сравнительно новые приборы, и поэтому неудивительно, что их возможности используются еще далеко не полностью. Однако возможно, что в ближайшем будущем мы увидим новые оригинальные случаи применения солнечных батарей.

ЗАДАЧИ

1. Точечный фототранзистор, характеристики которого показаны на рис. 8-11, включен последовательно с батареей 60 *в* и реле сопротивлением 10 *Ком*. Реле срабатывает при постоянном токе 3 *ма*. Какой нужен для этого световой поток (в миллилюменах)?

2. Если в эквивалентной схеме фотогальванического элемента (рис. 8-17) пренебречь малым сопротивлением выпрямителя *r*, то, как нетрудно видеть, фотогальванический элемент становится подобным генератору, включенному параллельно с выпрямителем. Чувствительность селенового фотоэлемента равна 30 *ма/лм*, а выпрямляющая часть его имеет характеристики в прямом направлении, соответствующие показанным на рис. 5-11. Определите и постройте как функцию светового потока в пределах от нуля до 1 *люм* график тока, даваемого этим элементом при внешней нагрузке 50 *ом*.

3. Составьте схемы управляемого светом триггера, использующего фотоумножителей элемент с *n-p-n* переходом, характеристики последнего приведены на рис. 8-21. Опишите его работу и рассчитайте напряжение источника питания, нагрузку, ток в обоих устойчивых положениях, внутреннее сопротивление в этих положениях и световой поток, необходимый для срабатывания.

ЛИТЕРАТУРА

1. Zworykin V. K, Ramberg E. G, Photoelectricity, Wiley, New York, 1947.

2 Walker R. C., Photoelectric Cells in Industry, Pitman, New York, 1947.

3. Busch G. L., Photoelectric Transmitting Typewriter, Electronics, April 1947, vol. 20, p. 96

4. Lippel B., A High-Precision Analog-to-Digital Converter, Proceedings of the National Electronics Conference, 1951, vol. 7, p. 206.

5 Folligstad H. G, Shive J. N., Yaeger R. E, A Transistor Optical Position Encoder and Digit Register, Proceedings of the National Electronics Conference, 1952, vol. 8, p. 766

6. Kernahan J. J. J., A Digital Code Wheel, Bell Laboratories Record, 1954, vol. 32, p. 766.

7. Hampton L. N., Newsom J. B., The Card Translator for Nation-Wide Dialing, Bell System Technical Journal, 1953, vol. 32, p. 1037.
8. См. [Л. 1] этой главы, а также гл 15 и 16
9. См. [Л. 2] этой главы, а также гл. 10 и 11.
10. Cashman R. J., Lead Sulfide Photoconductive Cells for Sound Reproduction, Journal of the Society of Motion Picture Engineers, 1947, vol. 49, p. 342.
11. Лукьянов С Ю, Фотоэлементы, Изд. АН СССР, 1948.
12. Козырев Б. П., Фотоэлементы и их применение, Труды первой межвузовской конференции по современной технике диэлектриков и полупроводников, июнь 1956, Изд. Ленинградского электротехнического института имени В И Ульянова (Ленина), 1957, 225—245

ГЛАВА ДЕСЯТАЯ

ТРАНЗИСТОРЫ

10-1. ВВЕДЕНИЕ

В 1948 г. фирма Bell Telephone Laboratories сообщила о разработке прибора, который может выполнять самую важную функцию вакуумного триода—усиливать сигналы. Было решено назвать его транзистором. Приставка trans указывает на свойство прибора передавать сигнал, когда он работает в режиме усилителя, а корень istor обозначает, что этот полупроводниковый элемент принадлежит к тому же семейству, что и термисторы, варисторы и пр.

Транзистор имеет много существенных преимуществ перед электронной лампой. Для его работы не нужен вакуум [1]. Он не имеет нити накала и, значит, не потребляет энергии накала, не требует времени для разогрева. Он может работать при напряжениях питания всего лишь 1—2 *в*. Он меньше и легче любой из имеющихся в продаже электронных ламп; благодаря малым размерам транзистор обладает значительной большей механической прочностью, чем вакуумные приборы. Несмотря на ограничения по мощности и полосе частот, он способен выполнять почти все функции, выполняемые элек-

---

[1] К сожалению, требования к герметизации многих современных транзисторов не отличаются от требований для электронных ламп. *Прим. ред.*

тронными лампами. Его малые размеры и небольшая потребляемая им мощность привлекательны в тех случаях, когда электронные лампы и их цепи геометрически и экономически себя не оправдывают.

Идея создания твердого усилительного устройства опередила фактическое изобретение транзистора более чем на десятилетие. Бóльшая часть подобных идей была направлена на имитацию устройства вакуумного триода внутри твердого выпрямителя путем введения какого-либо управляющего элемента в область пространственного заряда запирающего слоя. Однако разместить управляющую сетку в слое толщиной порядка 10^{-4} *см* чрезвычайно трудно, и поэтому все такие попытки оказались безуспешными.

Разработка транзистора пошла по совсем иному направлению. Он был открыт в процессе теоретических и экспериментальных исследований влияния на сопротивление тонкого слоя полупроводника электрических полей, достаточно сильных, чтобы проникнуть в толщу слоев [Л. 1 и 2]. В процессе этих исследований Бардин, Браттейн и Шокли [Л. 3 и 4] обнаружили, что если два соответствующим образом изготовленных контакта расположены близко друг от друга на поверхности полупроводника, то электрическую характеристику между одним из этих контактов и общим электродом, расположенным где-либо в другом месте полупроводника, можно изменять и этими изменениями можно управлять путем пропускания тока через другой контакт. Степень этого воздействия оказалась такой, что удалось получить усиление по мощности. В 1956 г. изобретатели были удостоены Нобелевской премии по физике за это открытие и за работу, которую они провели впоследствии, заложив теоретический фундамент наших современных знаний о транзисторах.

Действие транзистора является следствием рассмотренных в гл. 2 двуполярных свойств полупроводников. Ток через полупроводник пропускается в основном в виде потока неосновных носителей. Величина этого тока определяется изменениями входного сигнала. Ток проходит через расположенный поблизости контакт, находящейся под таким напряжением и с такой нагрузкой в цепи, при которых становится возможным усиление мощности [Л. 5]. Не выглядит ли это подобно тому

что происходит с электронным потоком в электронной лампе? Безусловно, это именно так. Хотя в настоящее время некоторые моменты остаются все еще неясными, например, почему ток в транзисторе должен обусловливаться потоком неосновных носителей, тем не менее совершенно очевидно, что имеется соответствие между процессами, происходящими в электронной лампе и транзисторе. Можно предположить, что для каждого существующего ныне электровакуумного прибора можно разработать соответствующий твердый прибор. Многие убеждены в этом, и ряд таких аналоговых твердых приборов уже построен, начиная с фотоэлектронных умножителей и кончая кольцевыми счетчиками.

10-2. ТОЧЕЧНЫЕ ТРАНЗИСТОРЫ

Исторически первым был описан точечный транзистор. Точечные транзисторы развивались в нескольких различных направлениях: одни — для работы на высоких частотах, другие — для больших мощностей, третьи — для особых случаев использования при коммутации и т. д.

Точечный транзистор состоит из небольшой германиевой пластинки, к поверхности которой с одной стороны близко друг к другу прижаты два заостренных электрода [Л. 7]. Эти контакты вместе с их электрическими вводами называются эмиттером и коллектором. Они выполняют те же функции, что и катод и анод в электронной лампе. На противоположной стороне пластинки располагается плоскостной контакт, называемый базой. В одном из видов точечных транзисторов все эти элементы заключены в цилиндрическую гильзу. Внешний вид таких транзисторов показан на рис. 10-1, а поперечный разрез — на рис. 10-2.

При изготовлении такого прибора из германиевого слитка, обладающего высоким сопротивлением, вырезается пластинка толщиной 0,5 *мм*, которая шлифуется с обеих сторон, меднится и лудится с одной стороны, а затем нарезается с помощью алмазной пилы на мелкие квадратики со стороной примерно 1 *мм*. При изготовлении транзистора один из таких квадратиков припаивается к латунному базовому выводу, который после травления поверхности германия запрессовывается в конец цилиндрической гильзы.

Эмитерный и коллекторный контакты изготовляются посредством запрессовки двух держателей в цилиндрическую фенопластовую втулку. Контактные проволочки эмиттера и коллектора толщиной около 0,1 *мм* изготов-

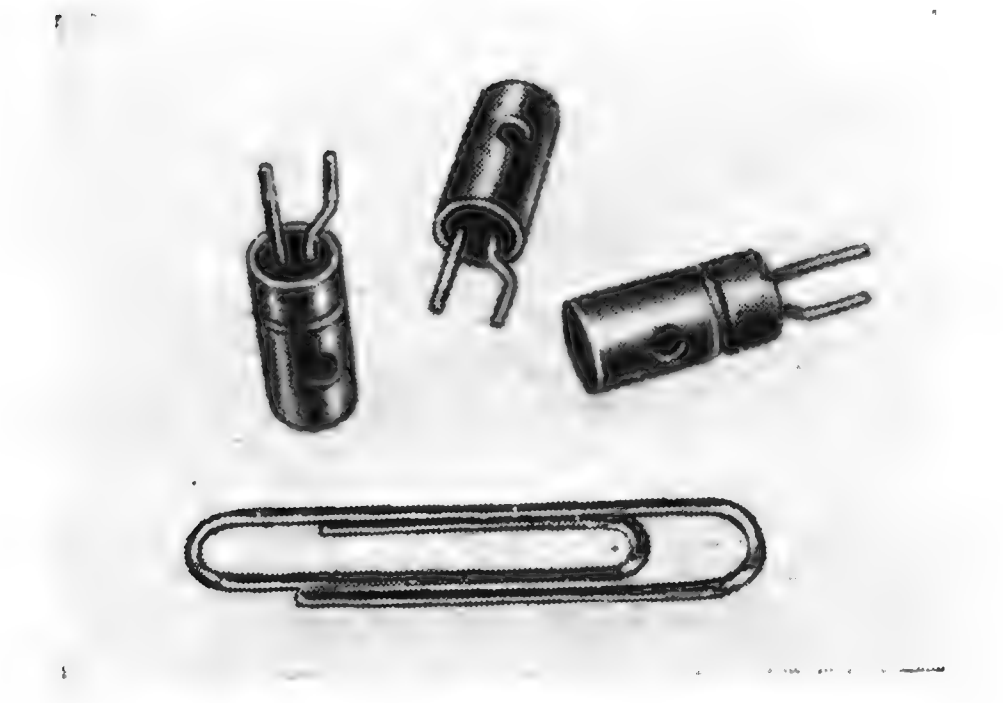


Рис 10-1 Внешний вид и размеры точечных транзисторов в металлических гильзах.

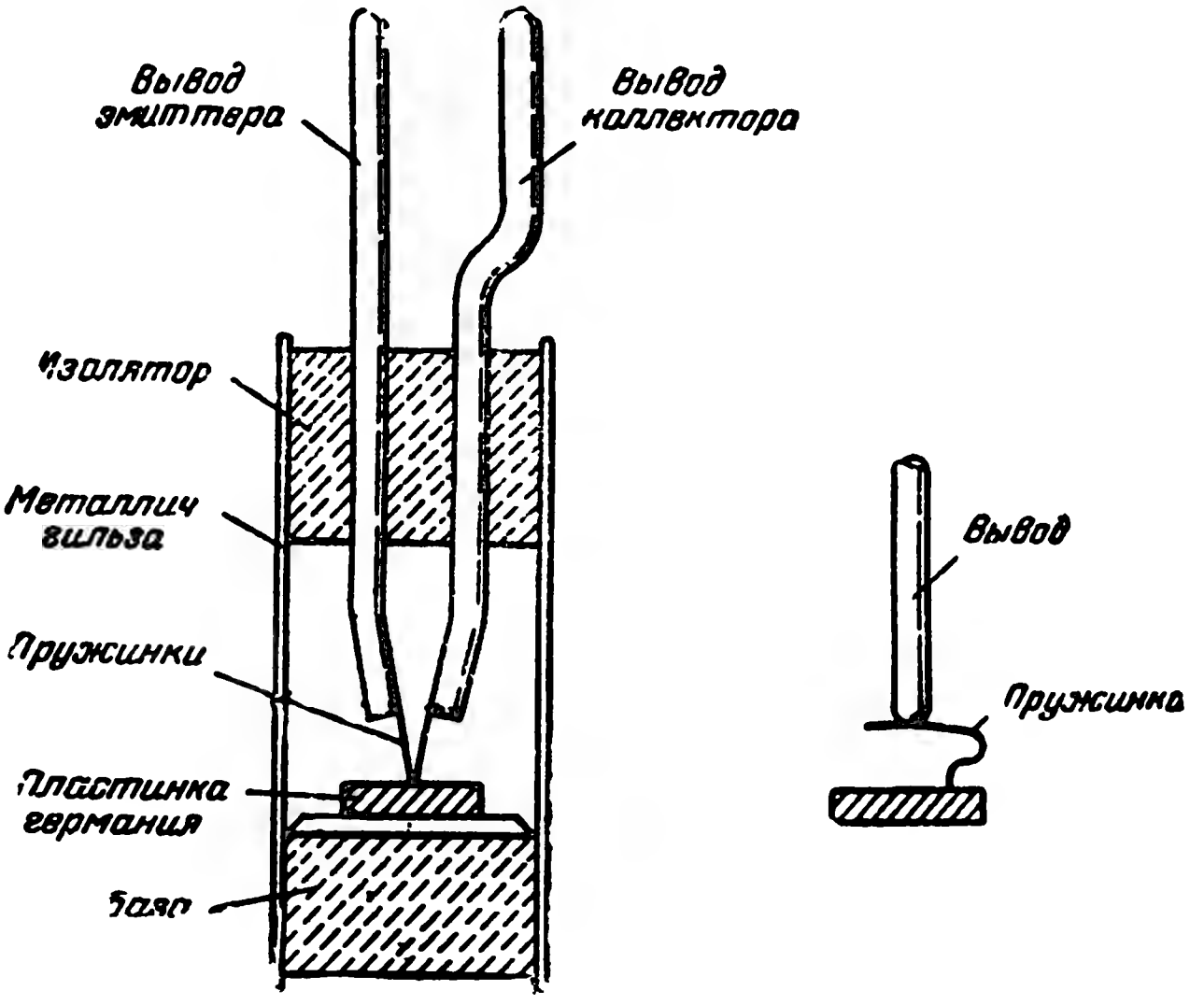


Рис. 10-2. Разрез точечного транзистора в металлической гильзе Видно расположение двух выпрямляющих контактов на поверхности полупроводника. Справа показан вид электродов сбоку.

ляют соответственно из бериллиевой и фосфористой бронзы, затачивают в форме лопаточек на концах, полируют и, наконец, припаивают к концам держателей. После этого контактные проволочки с помощью шаблонов превращают в плоские пружинки по форме, показанной на рис. 10-3. После окончательного подгибания проволочек для получения одинакового положения концов их в вертикальном направлении и требуемого расстояния между проволочками (0,025—0,05 *мм*) все контактное устройство вводится в гильзу до соприкосновения концов проволочек с поверхностью германия. Затем оно вдвигается в гильзу еще на 0,05 *мм*, чтобы создать некоторое нажатие пружинок и обеспечить надежность контактов при механических толчках, иногда происходящих при работе, а также в случае единичных перегрузок. При вертикальном нажатии на пружинки описанной формы не возникает касательных усилий, которые могли бы вызвать скольжение острий контактных проволочек по поверхности германия. Для повышения механической прочности и защиты поверхности германия от влаги внутренность гильзы заполняется заливочной массой. Выводы коллектора и эмиттера различаются по изгибу коллекторного вывода, как показано на рис. 10-2. Выводом базы служит металлической корпус.

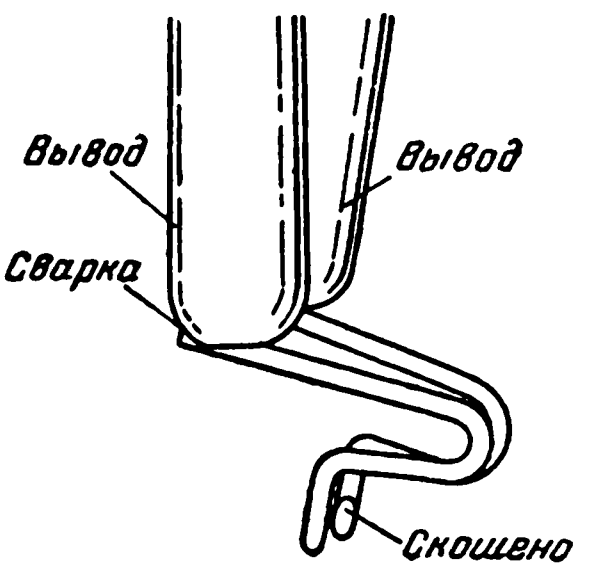


Рис 10-3 Форма эмитерной (бериллиевая бронза, диаметром 0,1 *мм*) и коллекторной (фосфористая бронза, диаметром 0,1 *мм*) контактных пружинок.

Смонтированные в гильзах, подобные описанным, транзисторы выпускались серийно, но недостаточная надежность их при работе во влажной среде, от которой заливка смолой не дает полной защиты, привела к разработке точечных транзисторов в герметически запаянных корпусах (рис. 10-4). В этом приборе геометрия контактов остается такой же, как и в транзисторах гильзового типа, и электрические характеристики их примерно одинаковы. Электроды прибора монтируются на трехштырьковом цоколе, а сверху надевается металли-

ческая оболочка, которая приваривается вдоль фланца. Внутренность оболочки заполняется инертным газом при атмосферном давлении.

Еще одна типичная конструкция точечного транзистора носит название двусторонней [Л. 8] или коаксиальной [Л 9]. В этом транзисторе эмитерный и коллекторный электроды расположены с противоположных сторон тонкой пластинки германия, с которой где-либо в другом месте обеспечивается базовый контакт. У одного из транзисторов такого вида, показанного на рис. 10-5, германий вырезан в виде цилиндрической шайбы, на которой с одной стороны вышлифовано углубление, так что толщина шайбы в центре составляет всего 0,05—0,075 *мм*. После травления шайба закрепляется внутри металлического кольцевого держателя и запрессовывается в средней части оболочки. Затем запрессовываются на место эмитерный и коллекторный контактные узлы так, что соответствующие контакты металлических проволочек с полупроводником приходятся друг против друга в центре вышлифованного углубления. Сборка заканчивается заливкой оболочки смолой.

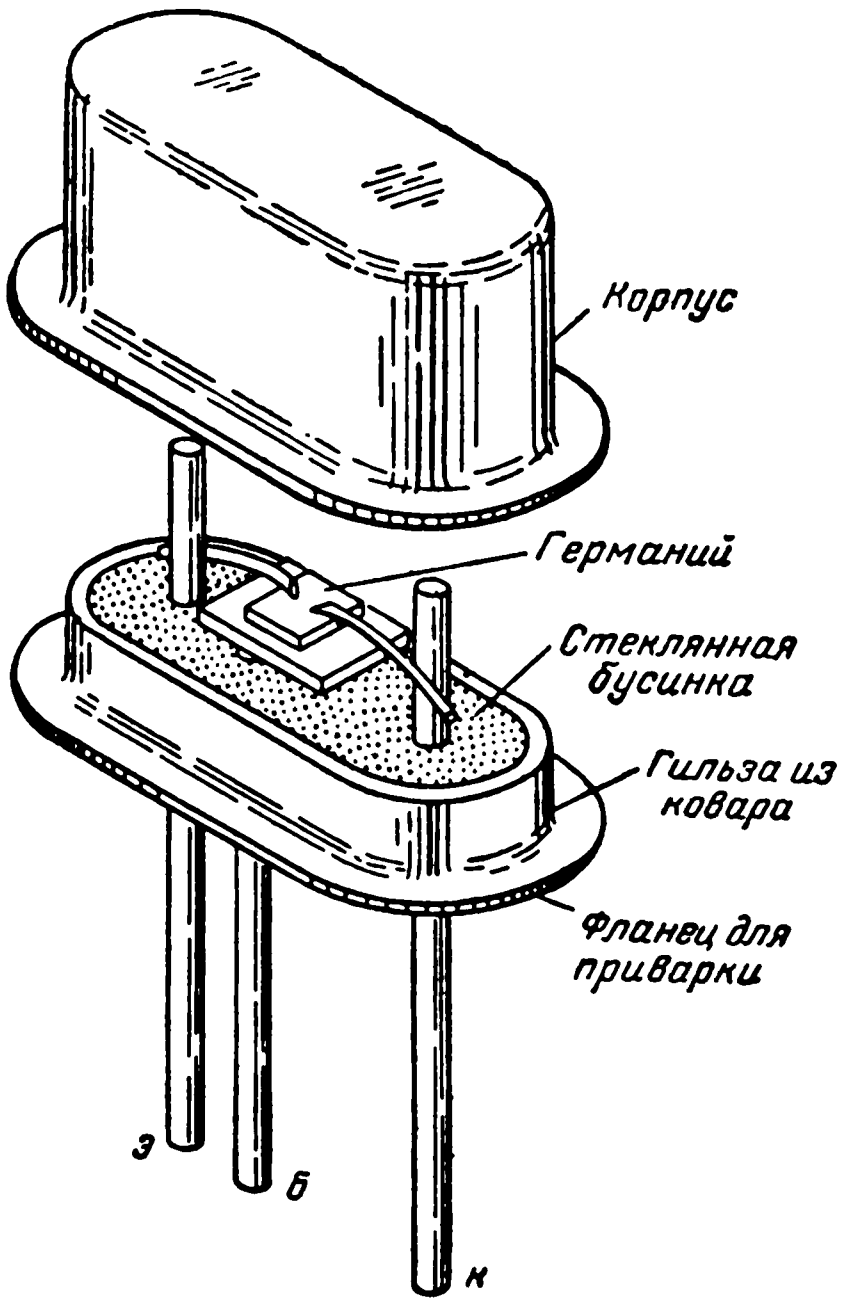


Рис 10-4 Точечный транзистор в корпусе из листового металла

Двусторонний транзистор сыграл полезную роль, впервые вполне определенно показав, что физические процессы в транзисторе происходят в объеме полупроводника, а не в поверхностном его слое, как это ранее предполагалось. Таким образом, этот транзистор помог

заложить некоторый экспериментальный фундамент, на котором Шокли и др. построили затем теорию плоскостных транзисторов, с помощью которой свойства и возможности этих приборов были предсказаны примерно за 2 года до того, как сами приборы были изготовлены и их параметры измерены.

Хотя после сборки всех трех описанных видов точечных транзисторов их можно считать механически за-

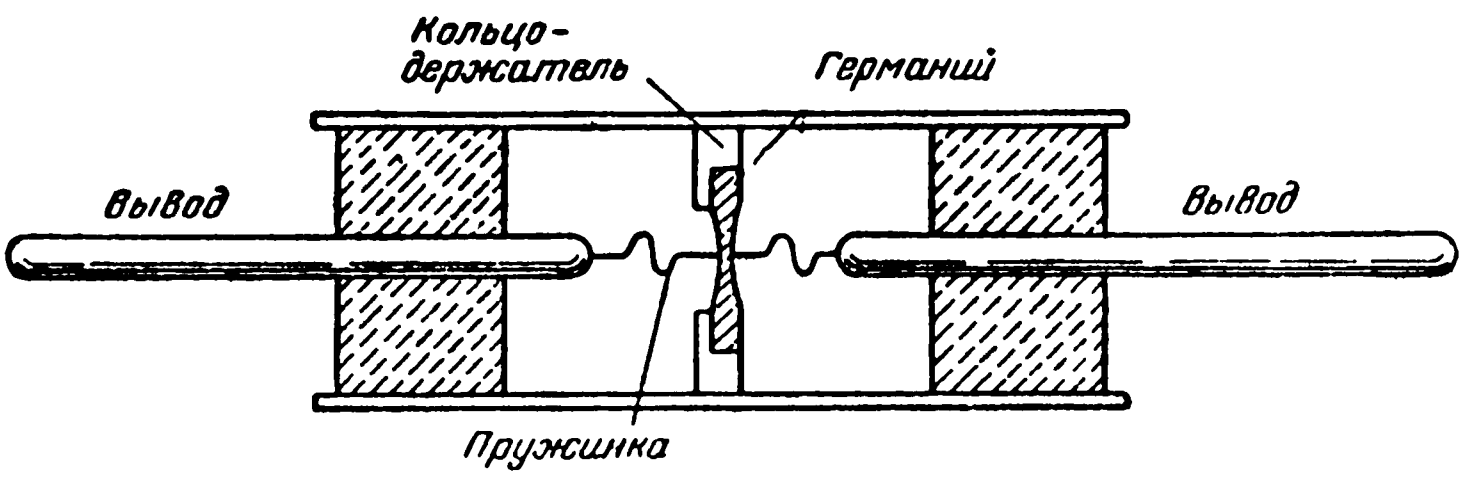


Рис 10-5. Коаксиальный точечный транзистор с эмитерно-коллекторными контактами на противоположных сторонах диска из полупроводника

конченными, они нуждаются в дальнейшей электрической обработке для получения хороших характеристик. Эта обработка называется тренировкой (формованием). Она заключается в пропускании импульсов тока через контакт коллектора в обратном направлении. Эти импульсы, обычно создаваемые разрядами конденсатора в цепи коллектора, постепенно усиливаются, пока не получаются желаемые характеристики, наблюдаемые на экране осциллографа после каждого импульса. Во время этого процесса в цепи эмиттера обычно поддерживается прямой ток. Сущность процессов, происходящих при тренировке (формовании), все еще неясна, но известно, что при этом происходит значительное улучшение характеристик уже готового транзистора [Л. 10].

10-3. ХАРАКТЕРИСТИКИ ТОЧЕЧНЫХ ТРАНЗИСТОРОВ

В цепь усилителя с общей базой транзистор включается, как показано на рис. 10-6. Указанные полярности напряжений соответствуют транзистору, изготовленному из полупроводника *n*-типа. На эмиттер подается небольшое прямое смещение постоянного тока

(несколько десятых долей вольта), на которое наложен усиливаемый переменный сигнал. На коллектор подано обратное смещение большей величины (единицы или десятки вольт) Сигнал, мощность которого вводится в полупроводник со стороны эмиттера, обладающего в режиме прямого тока низким сопротивлением, снимается с коллектора, обладающего высоким сопротивлением в режиме обратного напряжения; получаемое усиление по мощности равно отношению сопротивления коллектора к сопротивлению эмиттера, умноженному на квадрат коэффициента передачи тока. Такова элементарная схема процесса усиления с помощью транзистора

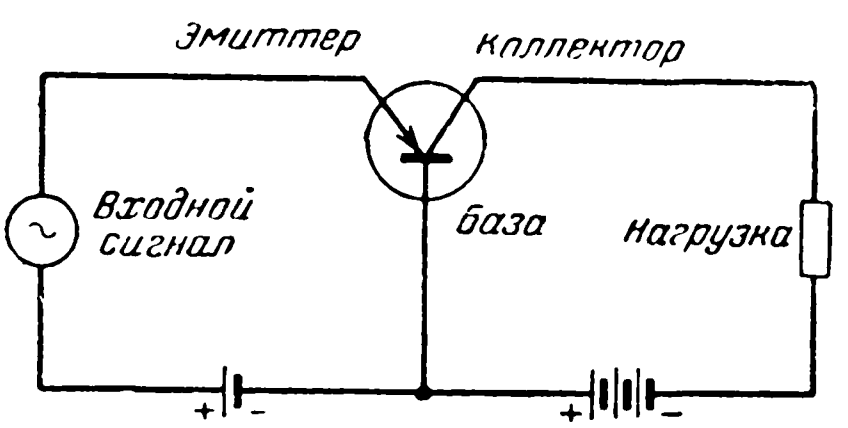


Рис. 10-6 Усилитель класса А на транзисторе по схеме с общей базой

На этой стадии изучения необходимо лишь добавить, что транзистор эффективно действует только при таких частях сигнала, которые передаются потоком *неосновных* носителей в полупроводнике.

Основные электрические характеристики транзисторов можно представить несколькими способами. Некоторые из них рассмотрены в следующей главе. Сейчас нам достаточно ограничиться описанием поведения транзисторов с помощью семейства выходных (коллекторных) характеристик при различных токах эмиттера. Семейство таких характеристик для типичного германиевого точечного транзистора *n*-типа показано на рис. 10-7.

Данные для построения такого семейства характеристик можно получить с помощью измерительной цепи для включения с общей базой, показанной на рис. 10-8. Изменяя напряжения питания на входных и выходных зажимах и одновременно считывая показания приборов $U_к$ и $I_к$ для нескольких комбинаций величин $U_к$ и $I_к$ при

постоянном токе эмиттера $I_э$, можно получить цифры, необходимые для построения одной из кривых, показанных на рис. 10-7. Подобным же образом, повторяя эту процедуру для других постоянных значений $I_э$, можно построить остальные кривые рассматриваемого семейства характеристик. Условимся считать ток положительным, если он втекает в транзистор через зажим, который

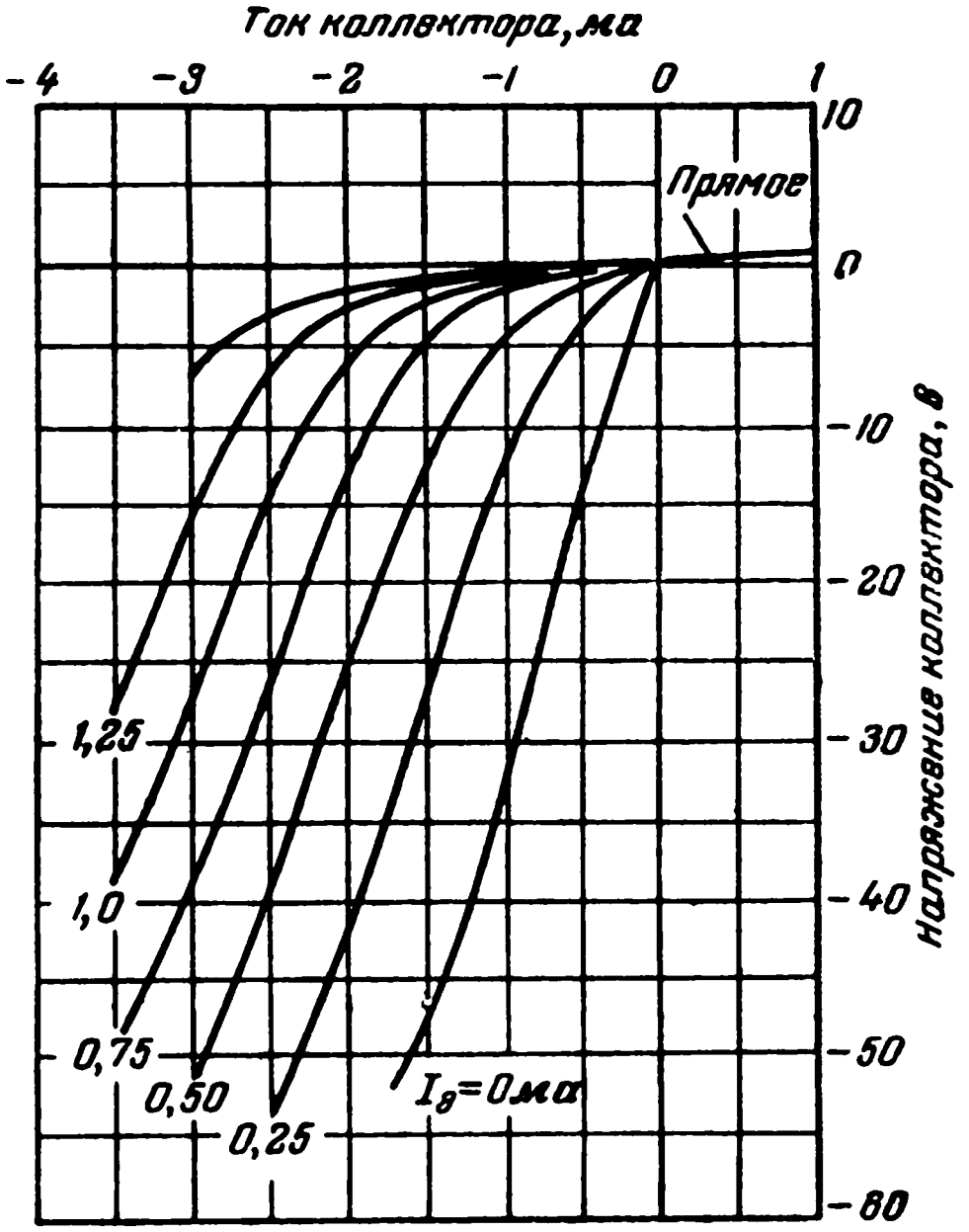


Рис. 10-7. Коллекторные характеристики точечного транзистора при разных токах эмиттера.

мы рассматриваем; напряжения измеряются по отношению к «земле», которой в случае, изображенном на рис. 10-8, является цепь базы.

Прежде всего заметим, что характеристики на рис. 10-7 располагаются в третьем квадранте. Это характерно для транзисторов, изготовленных из германия n-типа, какими является большинство точечных транзисторов. Кроме того, можно видеть, что в преде-

лах рабочих областей этих характеристик смещение коллектора отрицательно, что соответствует работе коллекторного контакта как выпрямителя при обратном напряжении. Одновременно в этих областях ток смещения эмиттера всегда положителен, что соответствует работе эмитерного контакта в режиме прямого напряжения.

Теперь становится яснее и физический смысл этого семейства кривых. Кривая $I_э=0$ представляет собой просто характеристику коллекторного перехода как выпрямителя в режиме обратного тока, снятую в то время, когда цепь эмиттера разомкнута и эмитерный ток не протекает Характеристику коллекторного перехода как

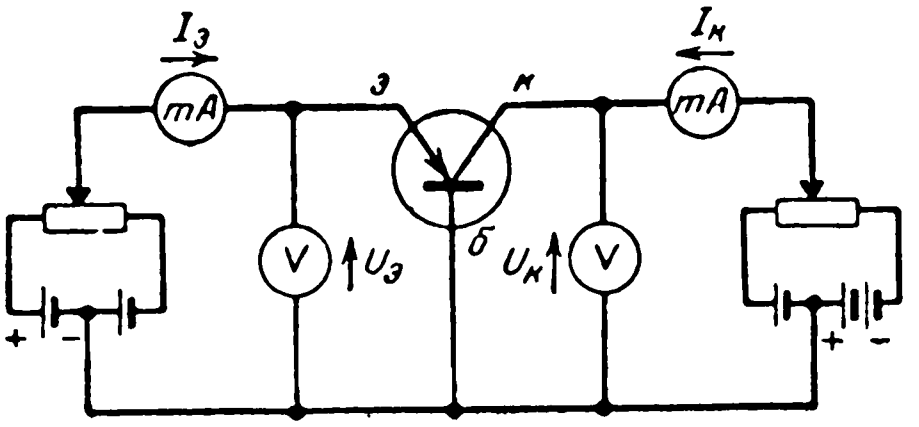


Рис 10-8 Экспериментальная цепь для снятия характеристик транзистора

выпрямителя в режиме прямого тока можно видеть в первом квадранте. Мы уже видели похожие характеристики у обыкновенного диода (гл 5). При замкнутой цепи эмиттера через транзистор проходит постоянный эмитерный ток и «диодные» коллекторные характеристики просто сдвигаются по направлению к бóльшим токам коллектора. Повторяя эту процедуру для разных постоянных токов эмиттера, можно построить все семейство кривых.

Существуют и другие способы графического представления характеристик транзистора. Так, например, можно построить семейство характеристик коллектора при постоянных значениях напряжения эмиттера вместо постоянных значений тока эмиттера; можно также представить эмитерные характеристики как функции тока коллектора; можно, наконец, показать семейство кривых тока коллектора как функции напряжения эмиттера для различных значений напряжений коллектора.

Мы выбрали систему координат, представленную на рис. 10-7, поскольку она содержит больше инженерного смысла, чем любое другое графическое описание свойств транзистора. По этим характеристикам очень просто определить такие величины, как полное сопротивление коллектора постоянному и переменному токам, необходимые величины напряжений и нагрузок и коэффициент передачи тока.

Читатель вправе задать вопрос: почему, представляя характеристики выпрямителя в гл. 5 с напряжением в качестве независимой величины, мы выбрали теперь для описания характеристик транзистора ток в качестве независимой переменной? Одной из причин этого является то, что физические процессы, происходящие при работе транзистора, легче описать рассматривая ток в качестве первопричины. Другое соображение связано с тем, что у некоторых характеристик эмиттера ток иногда оказывается двузначной функцией напряжения, тогда как напряжение всегда является однозначной функцией тока. В этом можно видеть существенное различие между транзисторами и электронными лампами. Электронная лампа обычно рассматривается как прибор, управляемый напряжением, тогда как транзистор является прибором, управляемым током. Электронная лампа стабильна в режиме короткого замыкания по переменному току, а транзистор в этом режиме нестабилен. Эти различия не заходят очень глубоко, но они показывают, что нельзя просто заменить электронные лампы в существующих схемах транзисторами.

Важное свойство точечного транзистора, заключается в том, что этот прибор усиливает переменный ток (рис. 10-7). Измерив расстояния между кривыми при постоянном напряжении коллектора, можно видеть, что в центральной области характеристик прирост тока коллектора значительно больше прироста тока эмиттера. Например, проведя через это семейство характеристик на уровне напряжения коллектора — 25 *в* горизонтальную нагрузочную линию, соответствующую короткому замыканию по переменному току, определим, что в точке, соответствущей току коллектора — 2 *ма*, изменение тока эмиттера на ±0,25 *ма* приводит к изменению тока коллектора на ±0,5 *ма*. Таким образом, прибор дает усиление по току, равное примерно 2. Этот коэф-

фициент передачи тока обозначается символом α и определяется как

$$\alpha = -\left.\frac{\partial I_к}{\partial I_э}\right|_{U_к=\text{const}}. \qquad (10\text{-}1)$$

Знак минус в уравнении (10-1) указывает, что согласно принятому обозначению направлений токи положительны, если они входят в транзистор из выводов эмиттера и коллектора.

При работе транзистора в режиме усилителя переменного тока усиление мощности частично обусловливается этим усилением тока, но в основном — большим различием полных сопротивлений входа (около 250 *ом*) и выхода (около 15—20 *ком*).

10-4. ЭЛЕМЕНТАРНАЯ ТЕОРИЯ МЕХАНИЗМА ДЕЙСТВИЯ ТРАНЗИСТОРА

Механизм действия транзистора был описан ранее в весьма элементарной форме как процесс, состоящий в перемещении неосновных носителей в полупроводнике.

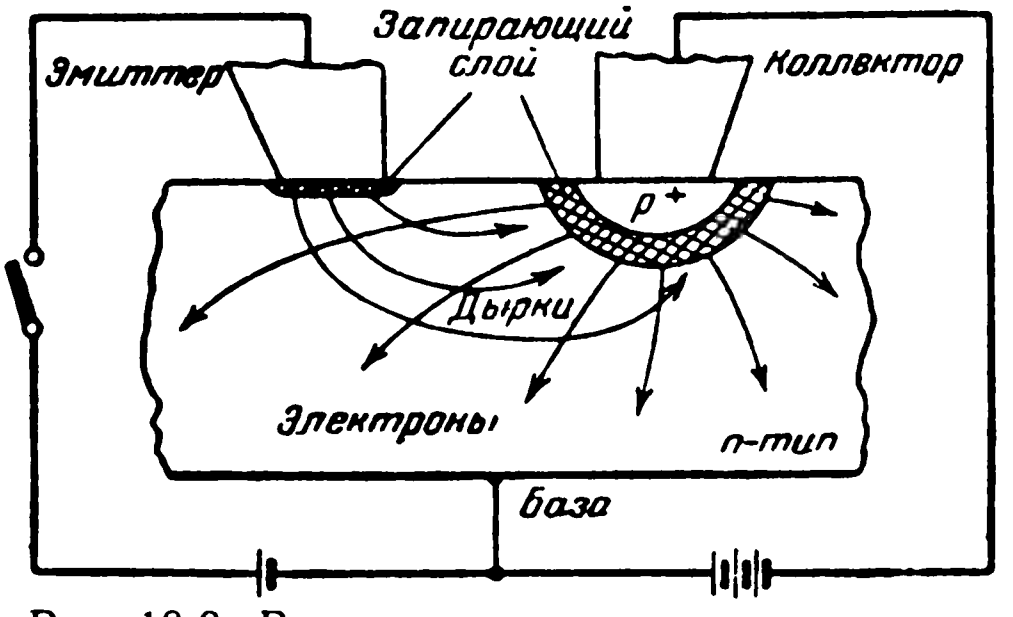


Рис. 10-9. В точечном транзисторе положительные дырки, вводимые в полупроводник из эмиттера, собираются на коллекторе, и в цепи его возникает ток.

Теперь мы опишем несколько более детально этот процесс, а в последних главах книги вернемся к нему, чтобы дать количественное описание процесса на более высоком уровне.

На рис. 10-9 показана в сильно увеличенном виде область вблизи эмитерного и коллекторного контактов точечного транзистора *n*-типа. Показана также связан-

ная с ней внешняя цепь, рассчитанная на использование прибора как усилителя при включении с общей базой. Выпрямительный слой эмиттера показан находящимся непосредственно под эмитерным контактом. Однако выпрямительный слой коллектора находится глубже в толще полупроводника, причем в этом слое содержится тонкая прослойка из материала *p*-типа. Конфигурация этого слоя определяется в процессе формования, которому подвергается коллектор. В течение этого процесса металл и полупроводник в точке контакта прогреваются в течение малых долей секунды до температуры в несколько сотен градусов. Предполагают, что в это время примеси, обусловливающие проводимость *p*-типа, диффундируют из металла коллектора в полупроводник, переводя его в *p*-тип в небольшой области германия, непосредственно примыкающей к коллекторному контакту [Л. 11] Это превращение происходит настолько интенсивно, что оно придает области *p*-типа почти металлическую проводимость, так что она начинает вести себя, как металлическое продолжение коллекторного контакта. Кроме того, в процессе формования коллекторный переход разрушается в том смысле, что он начинает пропускать от 10 до 100 раз больший обратный ток утечки, чем хороший точечный германиевый диод при этом же обратном напряжении. При разомкнутой цепи эмиттера этот ток утечки, представленный кривой $I_э=0$ на рис. 10-7, непрерывно протекает в виде электронного тока в полупроводнике *n*-типа от коллектора к базе и снова замыкается во внешней коллекторной цепи. Этот ток, расходящийся почти радиально от коллектора в толще полупроводника, создает на объемном сопротивлении полупроводника падение напряжения IR и слабое электрическое поле под эмитерным контактом, направленное непосредственно к коллектору.

Если теперь замкнуть цепь эмиттера, то через эмитерный выпрямляющий контакт начинает проходить прямой ток. Если этот контакт выполнен так, что бóльшая часть эмитерного тока состоит из положительных дырок, образованных на этом контакте и вводимых оттуда в полупроводник, то дырки будут отводиться этим полем к коллектору. Там они нейтрализуются, и ток, необходимый для этой нейтрализации, вызовет приращение тока обратной утечки, протекающего в цепи

коллектора. В хорошо изготовленном транзисторе n-типа ток эмиттера создается почти полностью дырками и почти все эти дырки можно собрать на коллекторе, за исключением пренебрежимо малой доли, потерянной по дороге из-за рекомбинации с электронами.

Выше мы предположили, что падение напряжения IR за счет обратного тока утечки коллектора образует поле между эмиттером и коллектором, которое в основном обусловливает перенос положительных дырок от эмиттера к коллектору. На самом деле это поле является лишь одной из причин их перемещения. Другой причиной является диффузия положительных дырок. Коллектор непрерывно отбирает эти дырки, и если у эмиттера образуются новые дырки, то возникает градиент концентрации дырок в пространстве между эмиттером и коллектором. Перемещение дырок от эмиттера к коллектору может в значительной части обусловливаться диффузией дырок по направлению этого градиента, даже если бы электрическое поле оказалось слишком слабым, чтобы быть эффективным. Действительно, до электрического формования коллектора наблюдается некоторое транзисторное действие, которое можно объяснить перемещением дырок путем диффузии. Предполагается, что формование настолько нарушает выпрямительное действие коллекторного перехода, что обратный ток утечки коллектора сильно возрастает и создает достаточно сильное электрическое поле между эмиттером и коллектором, которое помогает переносу дырок и, таким образом, увеличивает их сбор на коллекторе.

Этот механизм объясняет, как током коллектора может управлять ток эммитера, но он не объясняет, как данное изменение тока эмиттера может вызвать вдвое или втрое большее изменение тока коллектора. -Физическая сущность этого процесса умножения выяснена еще неполностью. Когда это явление наблюдалось впервые, было предположено, что поскольку положительные дырки проходят через весьма сильное электрическое поле у коллекторного перехода, они приобретают при этом достаточную энергию, чтобы создавать новые пары электронов и дырок за счет соударений. Это объяснение, однако, было оставлено, когда обнаружилось, что точечный транзистор проявляет

свойства умножения тока при столь низких напряжениях коллектора, при которых процесс умножения за счет соударений мало вероятен.

Теперь многие поддерживают другую гипотезу, заключающуюся в том, что поскольку в коллектор «втекает» ток дырок, из коллектора должен вытекать добавочный ток электронов для нейтрализации пространственного заряда приближающихся дырок. Так как в электрическом поле германия электроны движутся в 1,8 раз быстрее дырок, то при данном токе дырок, «входящих» в коллектор, для нейтрализации пространственного заряда у коллектора потребуется в 1,8 раза больший добавочный ток электронов, вытекающий из коллектора. Согласно этой модели общий ток коллектора должен быть представлен в виде суммы: 1) обратного тока утечки коллектора; 2) части тока эмиттера, обусловленной неосновными носителями, и, наконец, 3) добавочного тока нейтрализации пространственного заряда. В соответствии с этой моделью [Л. 12 и 13] коэффициент α должен иметь числовую величину около 2,8, что находится в хорошем согласии с величинами α, наблюдаемыми у серийных точечных транзисторов: в пределах от 2 до 3.

Еще одна модель, созданная для объяснения эффекта умножения тока: механизм «*n-p-n* ловушки» [Л. 14], основана на конфигурации эквипотенциалей в области и вблизи коллекторного перехода. Для понимания этой модели потребуется уяснить положения, которые развиты позже. Механизм *n-p-n* ловушки у коллектора может объяснить иногда встречающиеся аномально высокие величины α от 20 до 30.

10-5. ПЛОСКОСТНЫЕ ТРАНЗИСТОРЫ

Поскольку выпрямительное действие, обнаруживаемое металлическими контактами на полупроводниках, проявляется также и у *p-n* переходов в полупроводниках, то очевидно, по крайней мере теоретически, что можно получить транзисторное действие с помощью такого прибора, в котором эмитерный и коллекторный контакты заменены соответственно расположенными *p-n* переходами. Исторически возможности плоскостных транзисторов с *p-n* переходами были оценены [Л. 15] сперва теоретически, несколько раньше, чем тех-

нология получения таких переходов [Л. 16 и 17] была доведена до стадии практического изготовления плоскостных транзисторов. К настоящему времени технология получения переходов разработана до такой степени, что что *p-n* переход можно по желанию изготовлять в толще полупроводника или на его поверхности с помощью нескольких управляемых методов. Изготовленные таким образом плоскостные транзисторы мы и опишем. Пока же в качестве введения к изучению теории плоскостных транзисторов мы опишем два различных способа изготовления переходов: выращивание кристаллов с примесями и сплавление.

10-5-1. Транзисторы с выращенными переходами

Изготовление транзисторов с выращенными переходами начинается с приготовления слитка из кристаллического германия, имеющего тонкий (0,05 *мм*) слой германия *p*-типа, заключенный между областями *n*-типа, играющими роль контактов. Этот слиток изготовляется не сваркой трех кусков германия вместе (как можно было бы предположить), а выращиванием всего слитка в виде целого монокристалла с таким распределением примесей, которое дает желательное расположение областей проводимостей различных типов. Как это делается, описано в гл. 25.

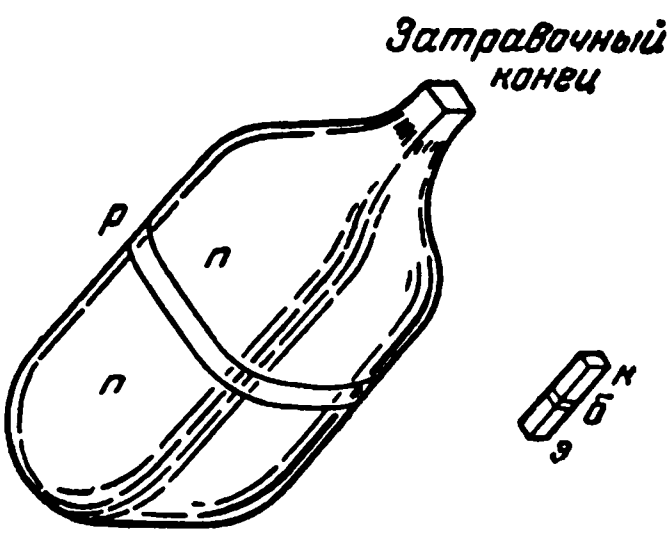


Рис. 10-10. Кристалл с выращенной *n-p-n* структурой, из которого можно вырезать много транзисторных элементов структуры *n-p-n*.

На рис. 10-10 показаны эскиз слитка типа *n-p-n*, а также эскиз небольшого элемента для транзистора, вырезанного из этого слитка. В типичном транзисторе со структурой *n-p-n* этот элемент может иметь длину около 3 *мм* и сечение около 1 *мм*^{2}. К этому элементу присоединяются невыпрямляющие выводные контакты путем припаивания проводов к концам элементов. К слою *p*-типа также присоединяется невыпрямляющий контакт привариванием заостренного конца тонкой проволочки к боковой стороне этого слоя. Механическое изготовление транзистора завершается заключением

n-p-n элемента в пластмассовую защитную оболочку или в герметически запаянный кожух для защиты от влаги.

На рис. 10-11 показан плоскостной транзистор типа *n-p-n*, работающий в качестве усилителя класса А по схеме с общей базой. Заметим, что хотя этот транзистор не точечной конструкции, он сохраняет основные черты структуры последней: эмиттер, базу и коллектор. Левый *p-n* переход выполняет те же функции, что и

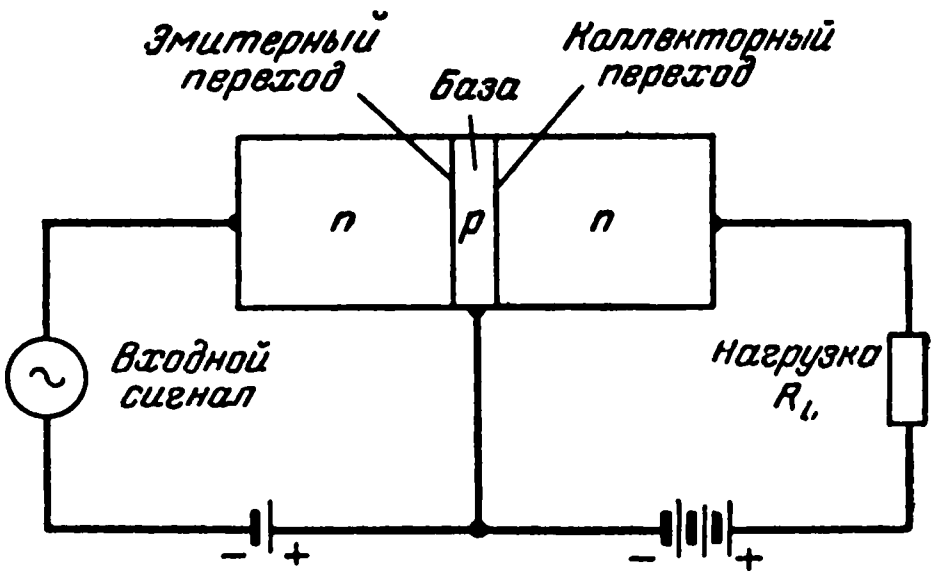


Рис. 10-11. Включение транзистора с *n-p-n* структурой в усилитель по схеме с общей базой.

эмитерный контакт в точечном транзисторе, а правый *p-n* переход выполняет функции коллекторного контакта. Тонкий базовый слой типа *p* в плоскостном приборе играет ту же роль, что полупроводниковая пластина в точечном приборе.

В плоскостном транзисторе входной сигнал переменной частоты подается на выпрямляющий слой между эмиттером и базой, находящийся под небольшим постоянным прямым смещением. В результате возникают большие изменения тока эмиттера при малых изменениях напряжения входного сигнала. Режим эмитерного перехода выбран так, что ток эмиттера состоит почти полностью из электронов, являющихся неосновными носителями в базовом слое *p*-типа, когда они попадают туда, пройдя через эмитерный переход. Попав в базовый слой, электроны диффундируют через него к коллекторному переходу, на который подается обратное смещение, чтобы собрать электроны. Таким образом, ток электронов прибавляется к обратному току утечки коллектора, который всегда течет в цепи коллектора и

базы. Поскольку коллекторный переход находится под обратным напряжением, он отбирает из базового слоя через область высокого сопротивления электронный ток, который ранее был введен в базу через область малого сопротивления. Поэтому усиление по мощности прибора фактически равно отношению этих сопротивлений.

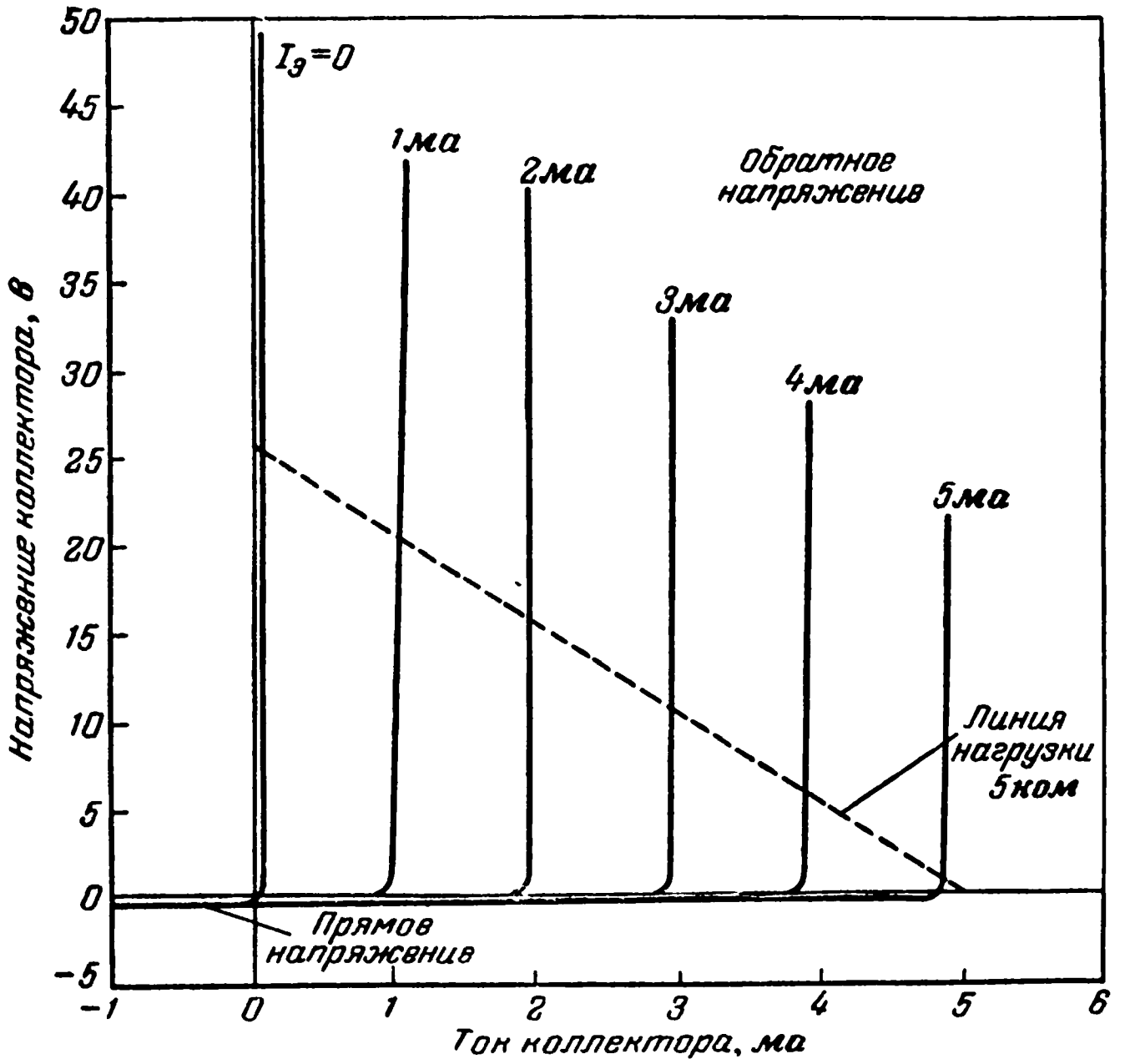


Рис. 10-12. Коллекторные характеристики транзистора с *n-p-n* структурой сильно отличаются от коллекторных характеристик точечного транзистора.

В приведенном выше описании можно узнать физические процессы, характерные для точечного транзистора. Их можно уяснить с помощью семейства характеристик на рис. 10-12. В транзисторе с *n-p-n* структурой на коллектор подается обратное напряжение, положительное по отношению к базе. Через коллекторный вывод транзистора протекает положительный ток; таким образом, коллекторные характеристики согласно принятым обозначениям полярности располагаются в первом квадранте прямоугольной системы координат. В пло-

скостном транзисторе коллектор не формуется электрически. При эмиттере, разомкнутом по постоянному току, характерны обратные токи насыщения коллектора всего в несколько микроампер и насыщение тока коллектора получается хорошим для любых токов эмиттера, вследствие чего семейство этих характеристик имеет вид семейства анодных характеристик вакуумного пентода. В плоскостных транзисторах отсутствует эффект умножения тока, поэтому величины α оказываются меньше единицы. Величина, на которую α в плоскостном транзисторе меньше единицы, определяется потерей небольшой части неосновных носителей из-за рекомбинации при диффузии их через базовый слой. Типичные величины α лежат в пределах от 0,95 до 0,995.

Заметим, что семейство характеристик занимает почти полностью всю площадь рабочего квадранта между осями тока и напряжения. Это означает, что, например, в схеме усилителя мгновенная рабочая точка может перемещаться в обе стороны по нагрузочной прямой почти от одной оси до другой. Такой режим соответствует феноменально высокому к. п. д. Для усилителей класса А на плоскостных транзисторах был получен к. п. д. 49%. Тщательные измерения коллекторных характеристик при весьма малых токах и напряжениях коллектора показывают, что «пентодообразная» форма кривых сохраняется при напряжениях коллектора всего 0,1 *в* и токах коллектора порядка нескольких микроампер. Это означает, что плоскостные транзисторы могут работать в качестве усилителей или генераторов при напряжении питания в десятые доли вольта и общем расходе мощности в несколько микроватт. В этом заключается одна из причин того, почему так интересуются транзисторами. Они могут выполнять различные функции в технике автоматики и связи при мощностях заметно ниже тех, к которым мы привыкли при работе с электромагнитными реле и электронными лампами.

10-5-2. Транзисторы со сплавными переходами

Выпрямляющие *p-n* переходы в полупроводниках могут быть изготовлены с помощью процесса сплавления [Л. 18, 19 и 20]. Если положить кусочек сурьмы размером с булавочную головку или свинца, содержащего сурьму в качестве примеси, на пластинку германия

p-типа и затем нагреть ее до температуры в несколько сотен градусов, то металл растворит некоторое количество германия и образует расплавленную капельку в образовавшемся углублении на поверхности германия. После охлаждения до комнатной температуры капелька металло-германиевого сплава затвердеет. В первую очередь затвердевает почти чистый германий, откладывающийся на кристаллической решетке исходной пластинки. Бóльшая часть металла постепенно выделяется при за-

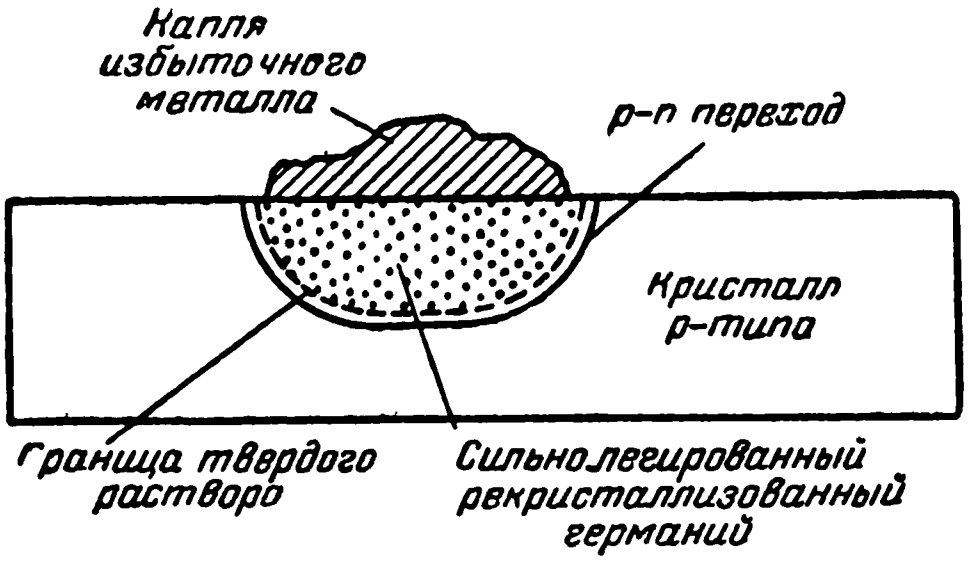


Рис. 10-13. Схема образования p-n перехода в пластинке германия при сплавлении.

стывании из раствора в виде пуговки, плотно припаянной к поверхности германия. Однако рекристаллизованный германий содержит в виде твердого раствора достаточное количество сурьмы, чтобы он приобрел резко выраженную проводимость n-типа. Таким образом, мы получим в пластинке p-типа включение n-типа с металлическим контактом, как показано на рис. 10-13.

В такой структуре сам p-n переход находится внутри основной пластинки несколько глубже границы между твердой и жидкой фазами, которая образуется в процессе расплавления. При повышенной температуре некоторые атомы сурьмы проникают в глубь германия за счет диффузии в твердом теле и, таким образом, образуют переход глубже границы рекристаллизации на несколько тысячных долей миллиметра. Именно поэтому переходы, изготовленные описанным выше способом, в ранних работах назывались «диффузионными переходами», хотя в настоящее время общепринято называть их «сплавными переходами».

Если исходную пластинку вырезать в определенной ориентации к осям кристалла, можно получить об-

ласть *n*-типа, имеющую не полусферическую форму, а почти совершенно плоскую. Если затем перевернуть эту пластинку и вплавить еще одну область *n*-типа с другой стороны пластинки против первой, то можно получить структуру типа *n-p-n*, подобную показанной на рис. 10-14. Припаяв подводящие проволочки к двум металлическим пуговкам и обеспечив в каком-нибудь месте металлический контакт с пластинкой *p*-типа, получим сплавной транзистор типа *n-p-n*. В производ-

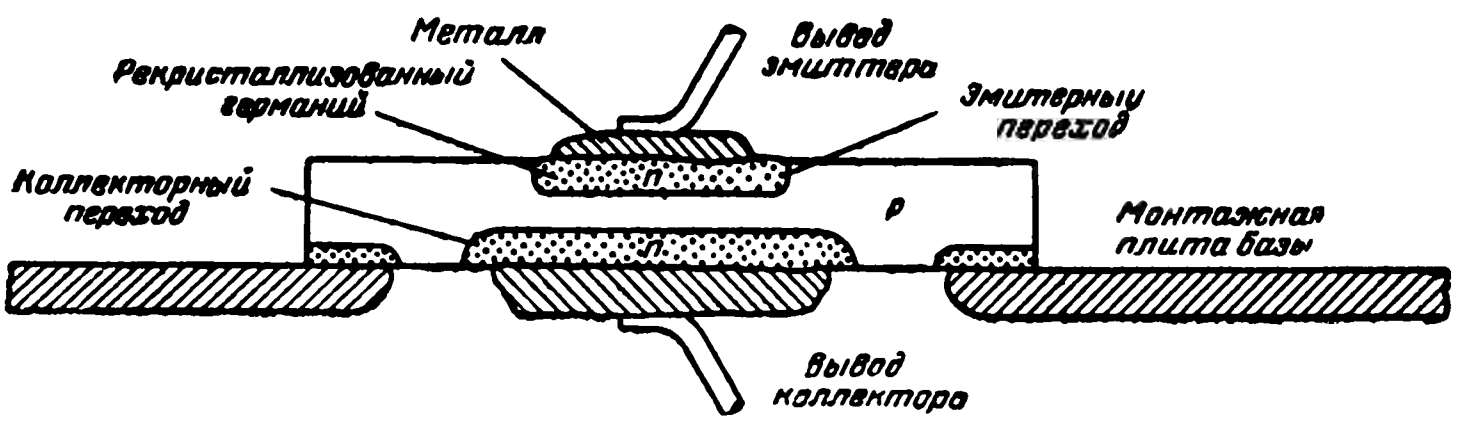


Рис. 10-14. Схема транзистора с *n-p-n* структурой, образованного двумя сплавными переходами на противоположных сторонах пластины полупроводника.

ственной практике исходная пластинка изготовляется в форме тонкого квадратика, который приплавляется краями к золотому держателю, имеющему круглое отверстие; в это отверстие помещается одна из сплавных пуговок. Все процессы сплавления выполняются в едином температурном цикле, причем вплавляемые металлические крупинки удерживаются на месте с помощью графитовых прижимов, пока она не вплавятся в германий. Затем базовая пластинка укрепляется на одном из штырьков стеклянно-металлического цоколя и заключается в герметический корпус, как показано на рис. 10-15. Основные размеры транзистора со сплавными переходами могут быть следующими: площадь германиевой пластинки — около 0,5 *см*^{2}, толщина германиевой пластинки 0,15 *мм*; диаметр сплавного участка *n*-типа[1] от 0,375 до 0,75 *мм*; глубина вплавления на каждой стороне от 0,05 до 0,075 *мм*.

Электрические характеристики транзисторов со сплавными переходами в основном подобны характеристикам

---

[1] Диаметр коллекторного вплавного электрода делается обычно больше эмитерного. Вследствие этого отвод неосновных носителей оказывается более полным, что приближает величины α к единице.

транзисторов с выращенными переходами сравнимой площади. Получаются «пентодообразные» семейства коллекторных характеристик с хорошо выраженным насыщением коллекторного тока. В процессе сплавления получаются переходы с толщиной запирающего слоя, меньшей, чем у выращенных переходов, поэтому напряжение пробоя коллектора для приборов со сплавными переходами ниже (около 50—60 *в* или менее) и

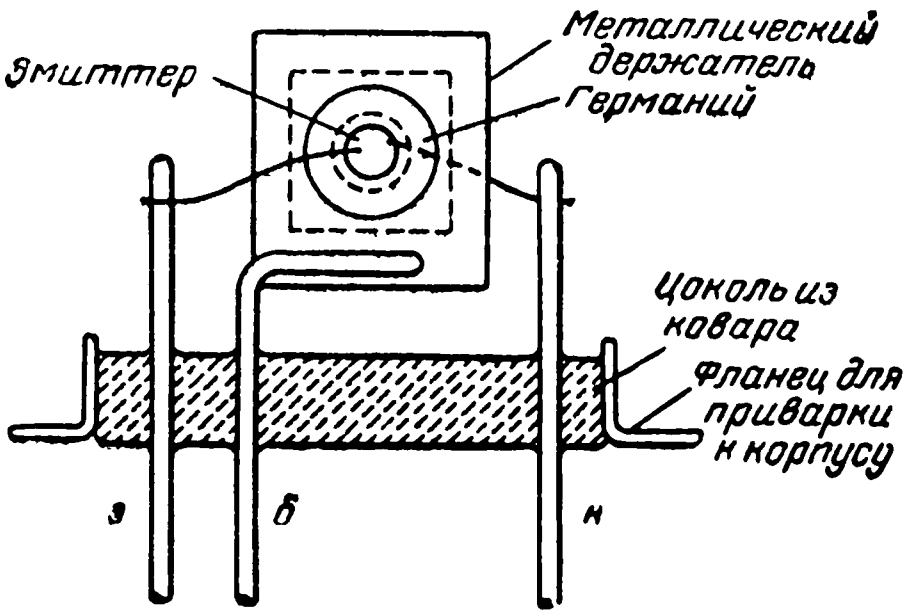


Рис. 10-15. Внутреннее устройство транзистора со сплавными переходами.

емкость переходов в 2—3 раза больше, чем у выращенных, при сравнимых площадях. Расстояние между эмитерным и коллекторным переходами в процессе сплавления можно по желанию довести до величины, меньшей 0,025 *мм*, если тщательно подготовить пластины, подобрать количества вещества во вплавляемых металлических каплях, выбрать время и температуру плавления.

При столь сходных характеристиках транзисторов со сплавными и выращенными переходами может создаться впечатление, что при их массовом использовании трудно отдать предпочтение какому-либо типу. В основном дело обстоит именно так; однако в некоторых случаях применения транзисторов, когда требуются какие-либо особые характеристики, можно сделать вполне определенный выбор. Так, например, если транзистор используется в триггерной ключевой схеме, часто бывает важно получить как можно меньшее сопротивление между эмиттером и коллектором. В транзисторе с выращенным переходом омическое сопротивление полу-

проводникового тела в эмитерной и коллекторной областях германиевого элемента добавляет несколько десятков ом к сопротивлению между эмиттером и коллектором, если прибор находится в проводящем состоянии и его сопротивление мало. В транзисторе со сплавными переходами рекристаллизованные области эмиттера и коллектора оказываются достаточно сильно легированными, так что их проводимость близка к металлической, а их сопротивление сквозному прохождению сигнала получается пренебрежимо малым. Поэтому в рассматриваемом случае предпочтительнее прибор сплавного типа. С другой стороны, для тех случаев, когда требуются высокие коллекторные напряжения, следует выбирать прибор с выращенным переходом.

Можно считать, что в настоящее время экономически выгоднее применять транзисторы со сплавными переходами. Изготовление приборов с выращенными переходами требует нескольких дорогостоящих индивидуальных операций. В производстве приборов со сливными переходами таких операций меньше и требуется менее тщательный индивидуальный контроль. Размещение вплавляемых капелек и процесс сплавления могут производиться одновременно на большом количестве изделий с помощью шаблонов. Будущее покажет, сохранятся ли эти качества у приборов обоих типов в процессе их дальнейшего совершенствования.

ЗАДАЧИ

1. Усиление по току $\left(\text{равное частной производной } \frac{dI_a}{dI_c}\right)$ обыкновенного электронного триода никогда не бывает больше единицы, а часто значительно меньше единицы из-за перехвата электронов сеткой и вторичной эмиссии с анода. Как в таком случае можно получить усиление по мощности, используя электронную лампу в схеме с общей сеткой?

2. Разберите возможности изготовления трехэлектродного жидкостного электролитического транзистора, в котором положительные и отрицательные ионы играют роль дырок и электронов. Изложите Ваши идеи примерно на одной странице. Каковы, по Вашему мнению, будут характеристики такого транзистора?

3. Представьте себе транзистор с плохим базовым контактом, который, отводя электроны из полупроводника n-типа, одновременно эмитирует в него некоторое количество дырок. Иначе говоря, базовый контакт сам по себе играет роль «частичного» эмиттера. Какие осложнения в работу транзистора будут внесены этим обстоятельством?

ЛИТЕРАТУРА

1. Shockley W., Pearson G. L., Modulation of Conductance of Thin Films of Semiconductors by Surface Charges, Physical Review, 1948, vol. 74, p. 232.

2. Bardeen J., Surface States and Rectification at a Metal-Semiconductor Contact, Physical Review, 1947, vol. 71, p. 717.

3. Bardeen J., Brattain W. H., The Transistor—A Semiconductor Triode, Physical Review, 1948, vol. 74, p. 230.

4. Brattain W. H., Bardeen J., Nature of the Forward Current in Germanium Point Contacts, Physical Review, vol. 74, p. 231.

5. Bardeen J., Brattain W. H., Physical Principles Involved in Transistor Action, Physical Review, 1949, vol. 75, p. 1208.

6. Marrows H. E., Transistor Engineering Reference Handbook, John F. Rider, Inc., New York, 1956.

7. Becker J. A., Shive J. N., The Transistor—A New Semiconductor Amplifier, Electrical Engineering, 1949, vol. 68, p. 215.

8. Shive J. N., The Double-Surface Transistor, Physical Riview, 1949, vol. 75, p. 69.

9. Kock W. E., Wallace R. L., The Coaxial Transistor, Electrical Engineering, 1949, vol. 68, p. 222.

10. Bardeen J., Pfann W. C., Effects of Electrical Forming on the Rectifying Barriers of n- and p-Germanium Transistors, Physical Review, 1950, vol. 77, p. 401.

11. Valdes L. V., Transistor Forming Effects in Germanium, Proceedings of the Institute of Radio Engineers, 1952, vol. 40, p. 445.

12. Hunter L. P., Physical Interpretation of Type A Transistor Characteristics, Physical Review, 1950, vol. 77, p. 558.

13. Sittner W. R., Current Multiplication in the Type A Transistor, Proceedings of the Institute of Radio Engineers, 1952, vol. 40, p. 448.

14. Shockley W., Electrons and Holes in Semiconductors, Van Nostrand, 1950 *; Theories of High Values of Alpha for Collector Contacts on Germanium, Physical Review, 1950, vol. 78, p. 294.

15. Shockley W., The Theory of p—n Junctions in Semiconductors and p—n Junction Transistors, Bell System Technical Journal, 1949, vol. 28, p. 435.

16. Shockley W., Sparks M., Teal G. K., p—n Junction Transistors, Physical Review, 1951, vol. 83, p. 151.

17. Wallace R. L., Pietenpol W. J., Some Circuit Properties and Applications of n—p—n Transistors, Bell System Technical Journal, 1951, vol. 30, p. 530.

18. Hall R. N., Dunlap W. C., p—n Junctions Prepared by Impurity Diffusion, Physical Review, 1950, vol. 80, p. 467.

19. Saby J. S., Fused Impurity p—n—p Junction Transistor, Proceedings of the Institute of Radio Engineers, 1952, vol. 40, p. 1358.

---

* См русский перевод: Шокли В., Теория электронных полупроводников, приложение к теории транзисторов, Изд иностранной литературы, 1953.

20. Ebers J. J., Alloyed Junction Transistor Development, Bell Laboratories Record, 1956, vol. 34, p. 8.
21. Полупроводниковые приборы и их применение, сборник статей под редакцией Я. А. Федотова, Изд. «Советское радио», вып. 1—7 с 1956 г.
22. Федотов Я. А. и Шмарцев Ю. М., Транзисторы, Изд. «Советское радио», 1960.

ГЛАВА ОДИННАДЦАТАЯ

ХАРАКТЕРИСТИКИ ТРАНЗИСТОРОВ И ЭЛЕМЕНТЫ СХЕМ

В предыдущей главе мы изучили транзисторы двух основных типов: точечные и плоскостные. Нам осталось еще рассмотреть некоторые менее очевидные различия между приборами этих типов не только в конструкциях, но и в характеристиках, механизме работы, усилительных и частотных свойствах, величинах внутренних сопротивлений и т. д. Ознакомимся сначала с тем, как описываются свойства транзистора, и с современной терминологией в этой области.

В настоящее время применяются три способа описаний свойств транзисторов. Первый из них — графические характеристики типа, описанного в предыдущей главе, о которых, однако, следует теперь сказать несколько подробнее. Второй способ основан на использовании эквивалентных схем. Транзистор представляется в виде эквивалентной цепи, состоящей из активных и пассивных элементов такого характера и так включенных, чтобы имитировать поведение транзистора. Если получается хорошее соответствие, то можно использовать эквивалентную схему в качестве основы для расчета поведения транзисторов при различных вариантах внешних цепей. Наконец, можно представить транзистор в виде четырехполюсника, заменяющего элемент цепи с четырьмя зажимами, внутреннее устройство которого для нас не играет роли, а внешние измеряемые характеристики можно выразить с помощью известных зависимостей между токами и напряжениями на зажимах. Каждый из этих трех способов имеет свои области применения.

11-1. ГРАФИЧЕСКИЕ ХАРАКТЕРИСТИКИ

Хотя транзистор имеет обычно всего три внешних вывода, он тем не менее представляет собой четырехполюсный прибор, если рассматривать его с точки зрения топологии цепей. В любой практической цепи транзистор имеет пару входных и пару выходных зажимов. Лишь случайно оказывается, что у транзисторов (как и у электронных ламп) один входной и один выходной зажимы физически и электрически объединены для любых схем включения прибора.

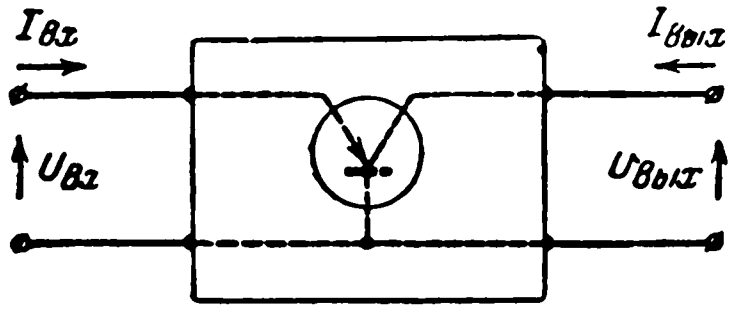


Рис. 11-1. Транзистор можно рассматривать как четырехполюсник с двумя входными и двумя выходными зажимами.

При рассмотрении транзистора в виде четырехполюсника [Л. 1], как показано на рис. 11-1, для полной характеристики прибора нужно исследовать четыре величины: входной ток $I_{вх}$, входное напряжение $U_{вх}$, выходной ток $I_{вых}$ и выходное напряжение $U_{вых}$. Эти величины связаны, и при всяком полном описании прибора данных должно быть достаточно, чтобы при известных значениях двух из этих величин можно было однозначно определить значения двух остальных величин. Так, например, два выражения, определяющие зависимости входного и выходного напряжений от входного и выходного токов, т. е.

$$\left.\begin{aligned} U_{вх} &= U_{вх}(I_{вх} I_{вых}); \\ U_{вых} &= U_{вых}(I_{вх} I_{вых}), \end{aligned}\right\} \qquad (11\text{-}1)$$

полностью характеризуют прибор независимо от того, заданы ли эти две функции аналитически или графически. Среди четырех упомянутых величин имеются другие пары зависимостей, которые обеспечивают равноценную полную характеристику четырехполюсника. Например, входной и выходной токи можно представить в виде аналитических или графических функций входного и выходного напряжений:

$$\left.\begin{aligned} I_{вх} &= I_{вх}(U_{вх} U_{вых}); \\ I_{вых} &= I_{вых}(U_{вх} U_{вых}). \end{aligned}\right\} \qquad (11\text{-}2)$$

Можно также ток и напряжение на входе или выходе выразить как функции тока или напряжения на других зажимах:

$$\left.\begin{aligned} I_{вх} &= I_{вх}(I_{вых} U_{вых}); \\ U_{вх} &= U_{вх}(I_{вых} U_{вых}); \end{aligned}\right\} \qquad (11\text{-}3)$$

$$\left.\begin{aligned} I_{вых} &= I_{вых}(I_{вх} U_{вх}) \\ U_{вых} &= U_{вых}(I_{вх} U_{вх}). \end{aligned}\right\} \qquad (11\text{-}4)$$

Наконец, можно написать перекрестно построенные пары уравнений, которые не повторяют ни одного из уравнений (11-1) — (11-4):

$$\left.\begin{aligned} I_{вх} &= I_{вх}(I_{вых} U_{вх}); \\ U_{вых} &= U_{вых}(I_{вых} U_{вх}); \end{aligned}\right\} \qquad (11\text{-}5)$$

$$\left.\begin{aligned} U_{вх} &= U_{вх}(I_{вх} U_{вых}); \\ I_{вых} &= I_{вых}(I_{вх} U_{вых}). \end{aligned}\right\} \qquad (11\text{-}6)$$

Любая из этих шести пар выражений (11-1) — (11-6) полностью характеризует четырехполюсник.

Рассмотрим теперь семейство коллекторых характеристик при различных значениях постоянного тока эмиттера, подобное представленному в предыдущей главе для точечного транзистора и повторно воспроизводимому для удобства на рис. 11-2,*а*. На нем представлено семейство кривых, описывающих зависимости между выходным (коллекторным) напряжением и входным (эмитерным) и выходным токами. Это семейство дает, таким образом, сведения в функциональной форме, содержащиеся во втором уравнении (11-1). Эти сведения необходимы и полезны, но они не дают полной характеристики. Для полноты характеристики необходимы дополнительные данные в форме, соответствующей первому уравнению (11-1), т. е. данные о зависимости напряжения эмиттера от токов эмиттера и коллектора. Эти дополнительные сведения представлены кривыми на рис. 11-2,*б*, где изображено семейство эмитерных характеристик для нескольких постоянных значений токов коллектора.

Два семейства характеристик на рис. 11-2,*а* и *б* полностью характеризуют транзистор. Любая точка в ра-

бочей области семейства коллекторных характеристик определяет $I_к$, $U_к$ и $I_э$. С этими данными можно перейти к семейству эмитерных характеристик и определить соответствующее $U_э$. Более того, эти два семейства обеспечивают возможность по значениям любых двух переменных определить значения двух других переменных.

Другие семейства характеристик, выражающие графически зависимости, соответствующие уравнениям

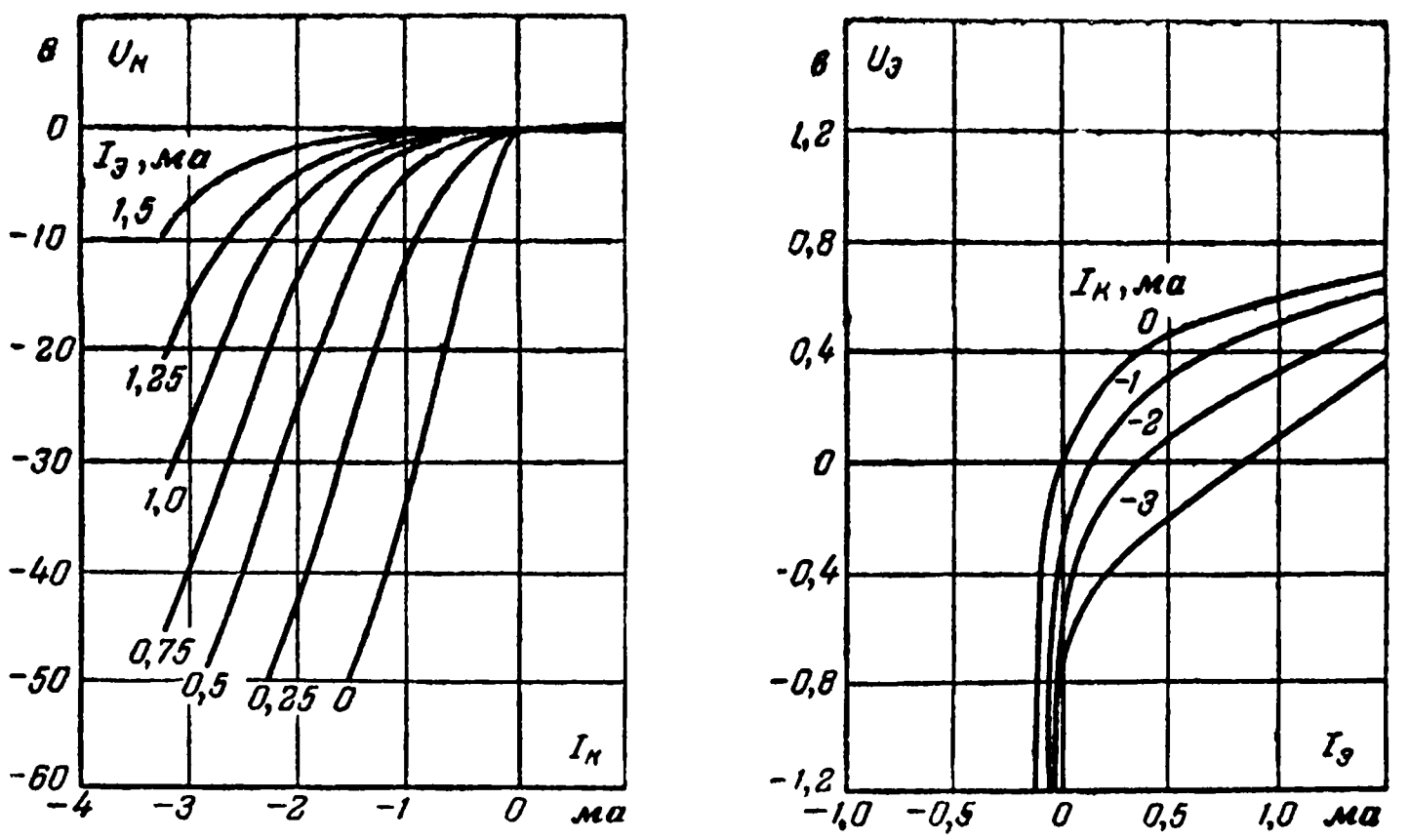


Рис. 11-2.

а—семейство коллекторных характеристик для различных токов эмиттера; *б*—семейство эмитерных характеристик для различных токов коллектора. Вместе эти семейства характеристик полностью характеризуют точечный транзистор.

(11-2)—(11-6), также могут быть использованы для полного описания прибора. Однако для инженерных расчетов особенно удобны и часто используются для описания поведения транзистора характеристики, соответствующие уравнению (11-1) и рис. 11-2. Между прочим, данные, представленные на рис. 11-2,*а*, могут быть графически описаны и другими способами, которые также хорошо показывают зависимость между напряжением коллектора и токами эмиттера и коллектора. Так, например, на семействе характеристик, показанном на рис. 11-3,*а*, представлены кривые напряжения коллектора как функции тока эмиттера для различных постоянных значений токов коллектора. Рис. 11-3,*а*, таким образом, дает те

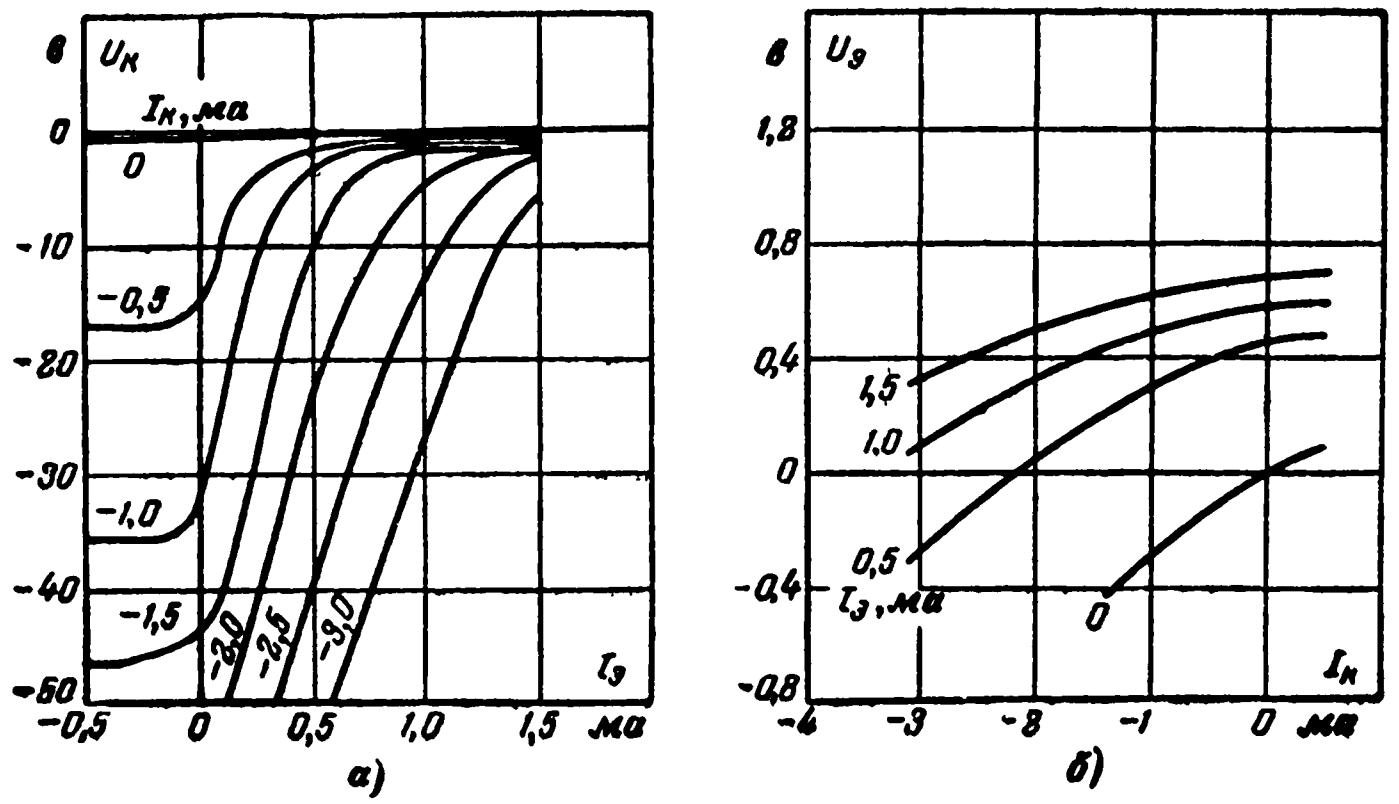


Рис. 11-3. Эти два семейства переходных характеристик дают те же сведения, что и семейства характеристик на рис 11-2, и также могут служить для описания точечного транзистора.

же сведения, что и рис. 11-2,*а*. Подобным же образом рис. 11-3,*б* дает те же сведения, что и рис. 11-2,*б*. Поэтому два семейства характеристик на рис. 11-3 дают другой способ полного описания поведения транзистора, основанный на уравнении (11-1).

11-2. ОПИСАНИЯ С ПОМОЩЬЮ ЭКВИВАЛЕНТНЫХ СХЕМ

Когда инженер рассматривает электронный прибор, он часто чувствует себя более уверенно, если может представить себе эквивалентную комбинацию схемных элементов, обладающих теми же внешними характеристиками, что и рассматриваемый прибор. Тогда, если он может каким-либо образом получить количественные данные о природе и поведения отдельных элементов, составляющих эту эквивалентную схему, он может использовать ее как основу для расчета поведения изображаемого ею прибора в различных цепях.

Существует способ представления электронного триода с помощью эквивалентной схемы переменного тока, показанной на рис. 11-4. В этой схеме r_c и r_a — сопротивления сетки и анода, а εe_a и μe_c — фиктивные генераторы напряжений, включение которых в цепь заставляет ее вести себя, как заменяемый триод. μ представляет собой коэффициент усиления лампы по напряжению. В типичном триоде ε практически пренебрежимо

мало. Далее, при работе на низкой частоте r_c практически бесконечно, хотя оно принимает конечную величину на более высоких частотах, когда начинают сказываться межэлектродные емкости сетка — анод и сетка — катод. Существенно отметить, что лампа представляет собой четырехполюсник и что для полной характеристики эквивалентной схемы требуется определение четырех величин: r_c, r_a, ε и μ.

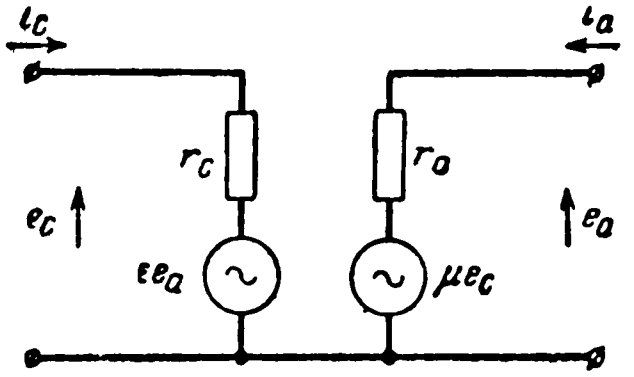


Рис. 11-4 При расчете электронная лампа может быть представлена такой эквивалентной схемой переменного тока.

Конечно, имеются и другие эквивалентные схемы электронного триода. Так, например, иногда используется схема, содержащая генераторы тока проводимости. Подобным же образом поведение транзистора может быть описано с помощью очень различных эквивалентных схем. Однако из всех этих схем две широко применяются в литературе по транзисторам, и каждый, кто изучает различные устройства на транзисторах, должен быть знаком с обеими. Эти две эквивалентные схемы показаны на рис. 11-5,*а* и *б*. Они выделены из всех возможных эквивалентных схем потому, что сочетания схемных элементов в них в значительной степени физически соответствуют структуре самого транзистора. В основе этих схем лежит простая Т-образная цепь, содержащая три сопротивления: $r_э$, $r_к$ и $r_б$, соответственно изображающие сопротивление переменному току эмитерного контакта или перехода, находящегося под прямым напряжением, сопротивление переменному току коллекторного перехода, находящегося под обратным напряжением, и сопротивление переменному току полупроводника между областью эмиттера и коллектора и базовым контактом. Эти три сопротивления соответственно называются сопротивлениями эмиттера, коллектора и базы.

Рассматриваемые две эквивалентные схемы различаются, однако, способами моделирования активного свойства транзистора, обусловливающего его усилительную способность. На схеме на рис. 11-5,*а* это свойство воспроизводится с помощью генератора напряжения,

включенного последовательно с сопротивлением коллектора. Прирост напряжения, даваемый этим генератором, численно равен приросту тока эмиттера $\Delta I_э$, умноженному на постоянную r_m, которая имеет физическую размерность сопротивления и является параметром данного транзистора. В другой эквивалентной схеме (рис. 11-5,*б*) активное свойство прибора обусловливается генератором тока, включенного параллельно с сопротивлением коллектора. Прирост тока генератора численно равен приросту тока эмиттера $\Delta I_э$, умноженному на безразмерную постоянную a, являющуюся параметром данного прибора. Разумеется, эти две схемы равноценны и для рассмотрения различных свойств транзистора можно использовать любую из них.

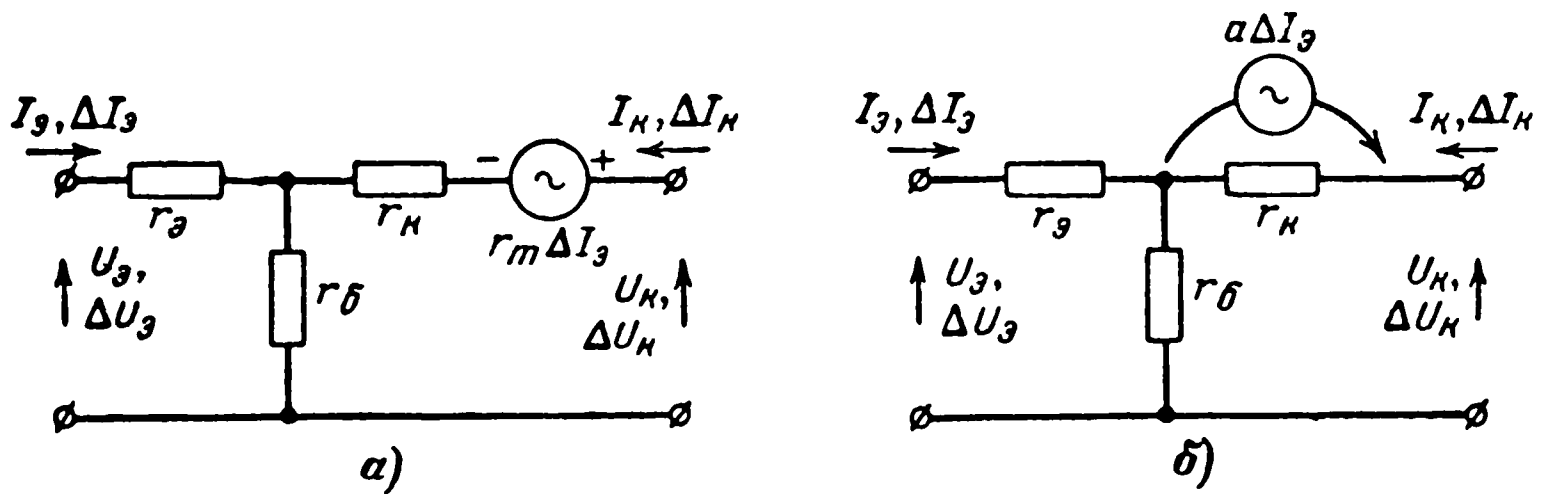


Рис 11-5 Эквивалентные схемы транзисторов
а — схема с эквивалентным генератором напряжения $r_m \Delta I_э$, *б* — схема с эквивалентным генератором тока $a\Delta I_э$.

Возможно, следует сделать несколько замечаний об эквивалентных генераторах тока. При изучении активных приборов, таких, как батареи и генераторы, обычно рассматривают их как идеальные генераторы напряжения, соединенные последовательно с внутренними сопротивлениями. Эта привычка возникла, вероятно, в результате обыкновения рассматривать напряжение как независимую переменную, а ток—как ее функцию, зависящую от нагрузки генератора напряжения. Однако имеется равноценный способ описания активного прибора как генератора *тока,* включенного *параллельно* его внутреннему сопротивлению. Ток, развиваемый таким эквивалентным генератором, является величиной постоянной, не зависящей от нагрузки, а ток, поступающий в нагрузку, зависит от того, как ток генератора распределяется между нагрузкой и внутренним со-

противлением генератора. Поэтому использование эквивалентного генератора тока для описания активных свойств транзистора в эквивалентной схеме, приведенной на рис. 11-5,*б*, так же оправдано, как и использование эквивалентного генератора напряжения в схеме на рис. 11-5,*а*. Однако первый способ несколько предпочтительнее, так как введение генератора тока более точно отображает физические процессы, происходящие в транзисторе.

В эквивалентной схеме на рис. 11-5,*б* коэффициент пропорциональности a можно назвать *внутренним* коэффициентом передачи тока. Его не следует смешивать с α—величиной, введенной в предыдущей главе для обозначения *внешнего* коэффициента передачи тока для постоянного напряжения коллектора. Однако a и α взаимозависимы. Если в эквивалентной схеме (рис. 11-5,*б*) зажимы коллектора и базы замкнуты между собой по переменному току накоротко (чтобы обеспечить постоянное напряжение коллектора), а на входе происходит прирост тока эмиттера $\Delta I_э$, то результирующий прирост тока коллектора будет:

$$-\Delta I_к = \Delta I_э \left(\frac{r_б}{r_б + r_к} + \frac{a r_к}{r_к + r_б} \right),$$

где знак минус обусловлен принятыми нами обозначениями входного и выходного токов.

Отсюда

$$\alpha = \lim_{\Delta I_э \to 0} \left(-\frac{\Delta I_к}{\Delta I_э} \right)_{U_к} = \left(-\frac{\partial I_к}{\partial I_э} \right)_{U_к} = \frac{r_б + a r_к}{r_б + r_к}, \qquad (11\text{-}7)$$

Уравнение (11-7) дает искомую зависимость между a и α. В реальных транзисторах точечного или плоскостного типа $r_к$ значительно больше $r_б$. Типичными величинами являются: для точечных транзисторов $r_к \approx$ $\approx 20\,000$ *ом*; $r_б \approx 150$ *ом*; для транзисторов с выращенным переходом $r_к \approx 5$ *Мом*; $r_б \approx 100$ *ом*. При таких значениях величин $r_к$ и $r_б$, как видно из уравнения (11-7), a и α численно почти равны. Это обстоятельство в прошлом привело к некоторой путанице в мыслях и выражениях, в результате чего внутренний генератор эквивалентной схемы иногда замещался величиной $\alpha \Delta I_э$. В нашей книге различие между a и α будет тщательно соблюдаться.

Значения параметров эквивалентной схемы $r_э$, $r_к$, $r_б$ и r_m или a (часто α) можно найти в паспортных данных различных типов и разновидностей транзисторов. Обычно указывают номинальные величины, иногда с верхними и нижними допусками. Кроме того, часто прилагаются графические характеристики. Для сравнения в табл. 11-1 приведены типичные величины параметров эквивалентных схем точечных и плоскостных транзисторов.

Таблица 11-1

Типичные значения параметров эквивалентных схем транзисторов

Параметр	Транзисторы		
	точечный	с выращенным *n-p-n* переходом	со сплавным *n-p-n* переходом
$r_э$, *ом*	100	50	25
$r_б$, *ом*	150	100	400
$r_к$, *ом*	20	5 000	1 000
a	2,5	0,99	0,99

Из сказанного выше очевидно, что рассмотренное описание транзистора с помощью эквивалентных схем характеризует поведение транзисторов в цепи переменного тока. Эквивалентные схемы являются схемами для переменного тока, и значения элементов эквивалентных схем определены при испытаниях транзисторов на переменном токе в определенных рабочих режимах, соответствующих заданному постоянному смещению. Иногда величины напряжения смещения указаны в паспортных данных; в противном случае величину напряжения смещения выбирают где-либо в середине площади, занимаемой семейством характеристик, в точке, где транзистор с наибольшей вероятностью будет использоваться в типичном режиме класса А.

Величины элементов эквивалентной схемы недоступны для непосредственного измерения; они должны быть вычислены из результатов измерений, произведенных во внешних цепях приборов. Как это делается, описано в § 11-4.

11-3. ТРАНЗИСТОР КАК ЧЕТЫРЕХПОЛЮСНИК

В § 11-1 мы рассмотрели способ описания транзистора как четырехполюсника с помощью одной из шести пар уравнений, определяющих взаимозависимости входных и выходных токов и напряжений. Такое описание можно дать для любого четырехполюсника, активного или пассивного, без какого-либо указания на его внутреннее содержание, произведя лишь необходимое количество надлежащих измерений на внешних выводах четырехполюсника. Мы рассмотрели также, как эти зависимости могут быть представлены графически в форме семейства характеристик.

Найдем теперь какой-нибудь простой способ аналитического выражения этих зависимостей для четырехполюсника. Например, предположим, что мы выбираем пару уравнений (11-1), описывающую взаимосвязь входных и выходных напряжений и токов:

$$\left.\begin{aligned} U_{\text{вх}} &= U_{\text{вх}}(I_{\text{вх}} I_{\text{вых}}); \\ U_{\text{вых}} &= U_{\text{вых}}(I_{\text{вх}} I_{\text{вых}}). \end{aligned}\right\} \qquad (11\text{-}1)$$

Предположим, что $I_{\text{вх}}$ и $I_{\text{вых}}$ обозначают величины постоянных составляющих входного и выходного токов в некоторой выбранной рабочей точке и что на них накладываются малые приращения токов $\Delta I_{\text{вх}}$ и $\Delta I_{\text{вых}}$. Тогда, если соответствующие приращения $\Delta U_{\text{вх}}$ и $\Delta U_{\text{вых}}$ входного и выходного напряжений являются линейными функциями приращений токов, мы можем написать:

$$\Delta U_{\text{вх}} = \frac{\partial U_{\text{вх}}}{\partial I_{\text{вх}}} \Delta I_{\text{вх}} + \frac{\partial U_{\text{вх}}}{\partial I_{\text{вых}}} \Delta I_{\text{вых}};$$

$$\Delta U_{\text{вых}} = \frac{\partial U_{\text{вых}}}{\partial I_{\text{вх}}} \Delta I_{\text{вх}} + \frac{\partial U_{\text{вых}}}{\partial I_{\text{вых}}} \Delta I_{\text{вых}}.$$

В этих уравнениях частные производные по одному какому-нибудь току указывают, что другой ток поддерживается постоянным, т. е. что соответствующая вторая пара зажимов поддерживается разомкнутой для переменного тока.

Можно принять, что эти дифференциальные выражения являются уравнениями для малых непрерывных изменений переменной составляющей тока. Введя для последних общепринятые обозначения для переменного тока, получим:

$$u_{вх} = \frac{\partial U_{вх}}{\partial I_{вх}} i_{вх} + \frac{\partial U_{вх}}{\partial I_{вх}} i_{вых};$$

$$u_{вых} = \frac{\partial U_{вых}}{\partial I_{вх}} i_{вх} + \frac{\partial U_{вх}}{\partial I_{вых}} i_{вых}.$$

В этой паре уравнений частные производные имеют физическую равномерность сопротивлений. В данном случае они имеют следующий смысл:

$$\left.\begin{aligned}
&\frac{\partial U_{вх}}{\partial I_{вх}} = z_{вх} && \text{— входное полное сопротивление для переменного тока при выходе, разомкнутом для переменного тока;}\\
&\frac{\partial U_{вх}}{\partial I_{вых}} = z_{о.прох} && \text{— обратное проходное полное сопротивление для переменного тока при входе, разомкнутом для переменного тока;}\\
&\frac{\partial U_{вых}}{dI_{вх}} = z_{п.прох} && \text{— прямое проходное полное сопротивление для переменного тока при выходе, разомкнутом для переменного тока;}\\
&\frac{\partial U_{вых}}{\partial I_{вых}} = z_{вых} && \text{— выходное полное сопротивление для переменного тока для входа, разомкнутого для переменного тока.}
\end{aligned}\right\} \quad (11\text{-}8)$$

Числовые значения четырех величин: $z_{вх}$, $z_{о\,прох}$, $z_{п\,прох}$ и $z_{вых}$ дают полную характеристику поведения транзистора в отношении малых сигналов переменного тока в рассматриваемой рабочей точке. Измерения этих четырех величин представляют собой несложную процедуру, при которой достаточно просто выполнить операции, соответствующие этим четырем частным производным. Так, для измерения $z_{вх}$ нужно разомкнуть выходные зажимы для переменного тока (например, введя дроссель последовательно с источником выходного напряжения) и измерить входное напряжение переменного тока, возникающее при пропускании переменного тока измеренной величины через входные зажимы.

Транзисторы часто характеризуются числовыми значениями четырех z, соответствующих уравнению (11-8), определенными в оговоренном или подразумеваемом режиме по постоянному току. В литературе по транзисто-

рам и паспортных данных прибора эти четыре z всегда обозначаются символами z_{11}, z_{12}, z_{21} и z_{22} соответственно, причем индекс 1 относится к входным зажимам, а индекс 2 — к выходным. В нашей книге предпочтение отдано индексам вх, вых и прох, чтобы уменьшить возможность затруднений с цифровыми индексами у читателей, впервые знакомящихся с материалом.

Если проделать такую же процедуру с одной из пяти других пар характеристических уравнений, то мы придем к другим семействам характеристических величин вместо четырех z, определяемых уравнением (11-8). Так, например, если мы для описания транзистора выберем выражения входных и выходных токов как функций входных и выходных напряжений согласно уравнению (11-2):

$$\left.\begin{aligned} I_{\text{вх}} &= I_{\text{вх}}(U_{\text{вх}}, U_{\text{вых}}); \\ I_{\text{вых}} &= I_{\text{вых}}(U_{\text{вх}}, U_{\text{вых}}), \end{aligned}\right\} \qquad (11\text{-}2)$$

то мы придем к паре таких характеристических выражений:

$$i_{\text{вх}} = \frac{\partial I_{\text{вх}}}{\partial U_{\text{вх}}} u_{\text{вх}} + \frac{\partial I_{\text{вх}}}{\partial U_{\text{вых}}} u_{\text{вых}};$$

$$i_{\text{вых}} = \frac{\partial I_{\text{вых}}}{\partial U_{\text{вх}}} u_{\text{вх}} + \frac{\partial I_{\text{вых}}}{\partial U_{\text{вых}}} u_{\text{вых}},$$

где
- $\frac{\partial I_{\text{вх}}}{\partial U_{\text{вх}}} = y_{\text{вх}}$ — входная проводимость для переменного тока для выхода, замкнутого для переменного тока;
- $\frac{\partial I_{\text{вх}}}{\partial U_{\text{вых}}} = y_{\text{о.прох}}$ — обратная проходная проводимость для переменного тока при входе, замкнутом для переменного тока;
- $\frac{\partial I_{\text{вых}}}{\partial U_{\text{вх}}} = y_{\text{п.прох}}$ — прямая проходная проводимость для переменного тока при выходе, замкнутом для переменного тока;
- $\frac{\partial I_{\text{вых}}}{\partial U_{\text{вых}}} = y_{\text{вых}}$ — выходная проводимость для переменного тока при входе, замкнутом для переменного тока.

(11-9)

Наше полное описание данного транзистора определяется числовыми значениями четырех проводимостей: $y_{вх}$, $y_{о\,прох}$, $y_{п\,прох}$, $y_{вых}$ в режиме, заданном постоянными смещениями. Такой способ описания можно найти в литературе по транзисторам, причем проводимости y всегда обозначаются как y_{11}, y_{12}, y_{21} и y_{22}.

В литературе о транзисторах встречается еще один способ описания их как четырехполюсника. Он основан на перекрестных зависимостях уравнений (11-6):

$$\left.\begin{aligned} U_{вх} &= U_{вх}(I_{вх}, U_{вых}); \\ I_{вых} &= I_{вых}(I_{вх}, U_{вых}). \end{aligned}\right\} \qquad (11\text{-}6)$$

После выкладок приходим к следующей паре характеристических выражений по переменному току:

$$u_{вх} = \frac{\partial U_{вх}}{\partial I_{вх}}\bigg|_{U_{вых}} i_{вх} + \frac{\partial U_{вх}}{\partial U_{вых}}\bigg|_{I_{вх}} u_{вых};$$

$$i_{вх} = \frac{\partial I_{вых}}{\partial I_{вх}}\bigg|_{U_{вых}} i_{вх} + \frac{\partial I_{вых}}{\partial U_{вых}}\bigg|_{I_{вх}} u_{вых}.$$

В этих выражениях все частные производные имеют различные физические размерности. Поэтому такой способ называется методом смешанных параметров, а частные производные обозначаются следующими символами:

$\frac{\partial U_{вх}}{\partial I_{вх}} = h_{вх}$ — входное полное сопротивление для переменного тока при выходе, замкнутом для переменного тока;

$\frac{\partial U_{вх}}{\partial U_{вых}} = h_{о.прох}$ — коэффициент обратной связи для переменного напряжения при входе, разомкнутом для переменного тока;

$\frac{\partial I_{вых}}{\partial I_{вх}} = h_{п.прох}$ — коэффициент передачи тока для переменного тока при выходе, замкнутом для переменного тока (величина, совпадающая с α согласно определению для транзистора в схеме с общей базой);

$\frac{\partial I_{вых}}{\partial U_{вых}} = h_{вых}$ — выходная полная проводимость для переменного тока при входе, разомкнутом для переменного тока.

Описание транзисторов с помощью значений четырех h — параметров часто встречается в литературе по транзисторам; параметры h обозначаются: h_{11}, h_{12}, h_{21} и h_{22}.

Выше приведены три различных способа описания характеристик транзистора, основанные на представлении транзистора как четырехполюсника. Каждый способ дает простое описание свойств транзистора по переменному току с помощью четырех величин z, либо четырех величин y, либо, наконец, четырех величин h. Эти величины можно довольно просто измерить на зажимах прибора. Выбор способа описания иногда определяется произвольно, а иногда диктуется практическим удобством применения имеющихся в распоряжении характеристических величин. Все три способа широко используются в литературе и паспортных данных по транзисторам.

11-4. ВЗАИМОСВЯЗЬ МЕЖДУ ТРЕМЯ СПОСОБАМИ ОПИСАНИЯ ТРАНЗИСТОРОВ

Итак, мы рассмотрели три основных способа описания свойств транзисторов: метод графических характеристик, метод эквивалентных схем и метод четырехполюсника (по параметрам z, y или h). Количественные значения всех этих характеристических параметров определяются с помощью измерений токов и напряжений на зажимах испытываемого прибора в оговоренных режимах питания и нагрузок.

Поскольку для описания одного и того же прибора можно использовать все три метода, очевидно, что все они должны быть взаимно эквивалентными. Иначе говоря, должны существовать некоторые определенные зависимости между характеристиками транзистора, значениями элементов его эквивалентной схемы и значениями его параметров как четырехполюсника. Практически, если имеется полное описание транзистора с помощью одного из методов, должен быть возможен переход к любому другому методу без необходимости дополнительных измерений Надобность в этом часто встречается Так, например, составитель схемы, имея в своем распоряжении семейства характеристик, приведенных в паспортных данных прибора, может захотеть определить из графиков числовые значения элементов эквивалентной схемы, чтобы воспользоваться ими при

расчете. Кроме того, мы еще не показали в § 11-3, как определить параметры элементов эквивалентной схемы на основании измерений на зажимах транзистора. Посмотрим теперь, как это можно сделать.

Предположим, что мы рассматриваем рис. 11-6, на котором приведены три способа описания характеристик транзисторов. Выберем для описания метод четырехполюсника по системе z-параметров и возьмем для полного графического описания рис. 11-2 (вместо рис. 11-3). Рассмотрим первый из z-параметров для метода четырех-

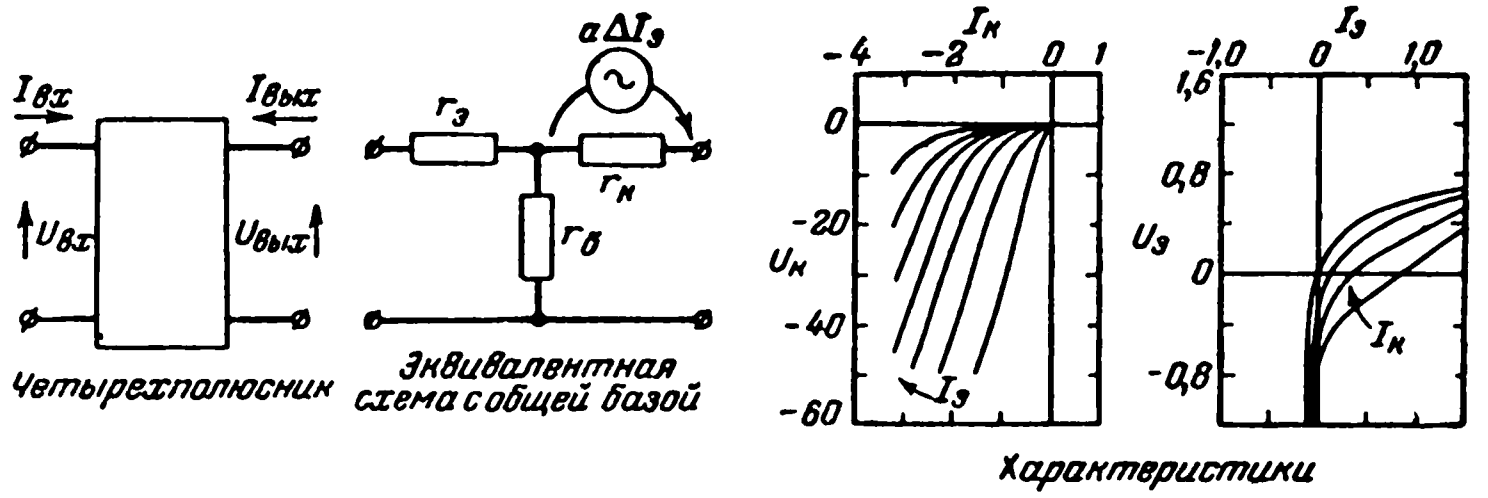


Рис 11-6 Несколько равноценных способов описания транзисторов.

полюсника: $z_{вх}$. Что он означает? По смыслу частной производной, которой он является, этот параметр представляет собой отношение приращения входного напряжения к приращению входного тока, которое его вызвало, при постоянном значении постоянной составляющей выходного тока (при выходе, разомкнутом для переменного тока). Если выполнить указанные операции для эквивалентной схемы, показанной в среднем столбце на рис. 11-6, то они приведут к полному сопротивлению $r_э + r_б$. Если же мы захотим узнать, что эта операция значит на языке графических характеристик (правый столбец), то мы заключим, что $z_{вх}$ численно равно крутизне (наклону) соответствующей кривой правого семейства в выбранной рабочей точке. Это и указывает на эквивалентность параметров при всех трех способах описания транзисторов.

Завершив это рассмотрение до конца, можно прийти к заключению об эквивалентности трех остальных z-параметров четырехполюсника, элементов эквивалентной схемы и наклонов и расстояний между кривыми графических характеристик. Эта эквивалентность сформулирована и представлена в следующей таблице.

Таблица 11-2

$z_{вх}=\frac{dU_{вх}}{dI_{вх}}$	$r_э + r_б$	Крутизна кривой на правом графике в рабочей точке, определенной величиной постоянного смещения
$z_{о.\,пр}=\frac{dU_{вх}}{dI_{вых}}$	$r_б$	Отношение расстояния между соседними кривыми на правом графике к соответствующему приращению тока коллектора при постоянном токе эмиттера (при перемещении по вертикали) в рабочей точке определенной постоянным смещением
$z_{п.пр}=\frac{dU_{вых}}{dI_{вх}}$	$r_б + ar_к$	Отношение расстояния между соседними кривыми на левом графике к соответствующему приращению тока эмиттера при постоянном токе коллектора (при перемещении по вертикали) в рабочей точке, определенной постоянным смещением
$z_{п.\,пр}=\frac{dU_{вых}}{dI_{вых}}$	$r_б + r_к$	Крутизна кривой на левом графике в рабочей точке, определенной постоянным смещением

Таким образом, показано, как параметры эквивалентной схемы $r_э$, $r_б$, $r_к$ и a могут быть определены на основании измерений четырех z-параметров в схеме четырехполюсника или из конфигурации графических характеристик.

Таким же способом можно прийти к эквивалентности параметров y или h для четырехполюсников и соответствующих параметров различных видов эквивалентных схем, а также различных возможных графических характеристик. Выполнение такой процедуры потребовало бы большой работы, но не оправдалось бы получением достаточного количества новых сведений о транзисторах. Кроме того, в задачи этой главы входит не исчерпывающее рассмотрение различных способов описания транзисторов, а лишь сообщение некоторых сведений об этом, чтобы читатель мог сознательно представить себе, что он найдет в литературе по этому вопросу. Поэтому мы ограничимся изложенным выше и перейдем к рассмотрению некоторых основных вопросов, касающихся цепей транзисторов.

11-5. ЦЕПИ ТРАНЗИСТОРОВ

В практической схеме усилителя с общей базой ко входу транзистора подключается генератор сигналов с внутренним сопротивлением $R_г$, а к выходу — сопротивление нагрузки $R_н$, на котором выделяется мощность усиливаемого сигнала. Такая цепь показана на рис. 11-7, где транзистор представлен в виде эквивалентной схемы для переменного тока. Элементы цепи постоянного тока на этой схеме не представлены, так как мы интересуемся только поведением усилителя при переменном токе и можем допустить, что источники напряжения смещения практически не имеют внутреннего сопротивления. Правда, режим постоянного тока в некоторой степени определяет при малых сигналах величины $r_э$, $r_б$, $r_к$ и a, но мы допустим, что рабочие точки эмиттера и коллектора уже выбраны и параметры эквивалентной схемы этого режима известны.

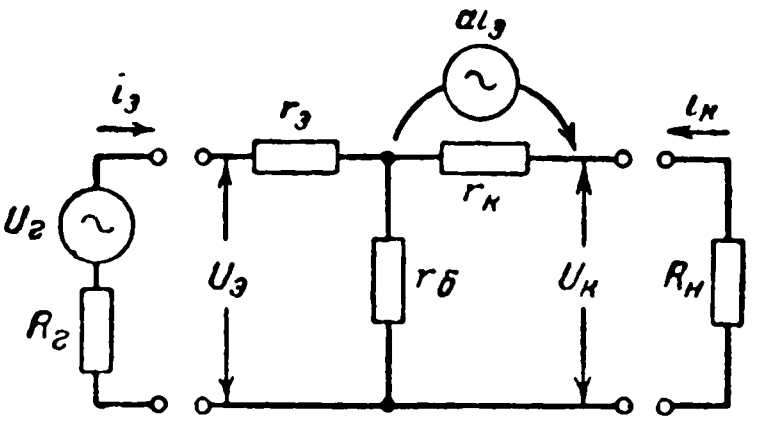


Рис. 11-7. Транзистор в усилителе по схеме с общей базой Токи и напряжения обозначены общепринятыми для переменного тока символами.

Изучим теперь работу этого транзистора в усилителе при заданном режиме. Каковы будут входное и выходное полные сопротивления схемы и какое усиление мощности она обеспечит? Как работа схемы зависит от величин сопротивлений генератора и нагрузки?

Не проводя детально все вычисления, укажем лишь на следующие результаты [Л. 2].

Входное полное сопротивление при включении с общей базой и выходной нагрузке $R_н$ равно:

$$r_э + r_б \left(\frac{r_к (1 - a) + R_н}{r_б + r_к + R_н} \right). \qquad (11\text{-}11)$$

Выходное полное сопротивление при включении с общей базой и входном сопротивлении генератора сигналов $R_г$ равно:

$$r_к + r_б \left(\frac{R_г + r_э - a r_к}{R_г + r_э + r_б} \right). \qquad (11\text{-}12)$$

Рабочий коэффициент усиления мощности между зажимами $R_г$ и $R_н$ при включении с общей базой в режиме малых сигналов равен:

$$4R_гR_н\left[\frac{r_б+ar_к}{(R_г+r_э+r_б)(r_к+r_б+R_н)-r_б(r_б+ar_к)}\right]^2. \quad (11\text{-}13)$$

Изучая эти выражения, мы обнаруживаем, что, как и можно было предполагать, входное полное сопротивление зависит только от параметров эквивалентной схемы транзистора и выходной нагрузки и вовсе не зависит от сопротивления входного генератора. Отметим также, что если $R_н$ беспредельно увеличивать, то входное сопротивление приближается к величине $r_э+r_б$, ожидаемой для режима цепи с разомкнутым выходом Подобным же образом выходное полное сопротивление зависит только от параметров эквивалентной схемы транзистора и сопротивления генератора и не зависит от сопротивления нагрузки. Если сопротивление генератора беспредельно увеличивать, то выходное полное сопротивление приближается к величине $r_к+r_б$, ожидаемой для режима цепи с разомкнутым входом.

Усиление мощности, определяемое из уравнения (11-13), зависит не только от параметров эквивалентной схемы транзистора, но и от обоих (входного и выходного) внешних сопротивлений: $R_г$ и $R_н$. Максимальное усиление мощности имеет место при одновременном согласовании обоих: $R_г$ — с входным полным сопротивлением транзистора, а $R_н$ — с выходным. Однако эти сопротивления зависят от того, что подключено к противоположной паре выводов, и экспериментальное определение величин $R_г$ и $R_н$ для их одновременного согласования требует подбора методом последовательных приближений. Аналитическое определение оптимальных входного и выходного полных сопротивлений усилителя с транзистором, включенным с общей базой, дает:

$$R_{г.опт}=(r_э+r_б)\sqrt{1-\frac{r_б(r_б+ar_к)}{(r_б+r_к)(r_э+r_б)}}; \quad (11\text{-}4)$$

$$R_{н.опт}=(r_к+r_б)\sqrt{1-\frac{r_б(r_б+ar_к)}{(r_б+r_к)(r_э+r_б)}}. \quad (11\text{-}15)$$

Максимальный коэффициент усиления мощности, обеспечиваемый при выполнении этих условий, равен:

$$\frac{1}{(r_б + r_к)(r_э + r_б)} \left[\frac{(r_б + ar_к)}{1 + \sqrt{1 \dfrac{r_б(r_б + ar_к)}{(r_б + r_к)(r_э + r_б)}}} \right]^2. \quad (11\text{-}16)$$

11-6. СХЕМЫ С ОБЩИМ ЭМИТТЕРОМ И ОБЩИМ КОЛЛЕКТОРОМ

До сих пор мы говорили о таком включении транзистора, когда база являлась общим электродом для входной и выходной частей схемы. Эквивалентная схема транзисторов, которую мы рассмотрели выше, также являлась схемой с общей базой, и выражения для сопротивлений и коэффициента усиления были выведены для транзисторов в схемах с общей базой. Включение транзистора с общей базой аналогично включению электронной лампы с общей сеткой при подведении сигнала к катоду. Такое использование электронной лампы несколько необычно, но оно совершенно реально, если входной сигнал способен «раскачать» сравнительно высокую емкость между катодом и землей.

Возможны два других включения транзистора (рис. 11-8). Это схемы включения транзистора с общим эмиттером и общим коллектором, соответствующие схемам включения электронных ламп с общим катодом и общим анодом (катодный повторитель).

Плоскостные транзисторы применяют обычно в схемах с общим эмиттером. При правильном согласовании входного и выходного полных сопротивлений такое включение позволяет получить большее усиление на каскад. Кроме того, при включении плоскостных транзисторов по схеме с общим эмиттером входное и выходное согласованные полные сопротивления меньше различаются, чем при включении с общей базой Это обстоятельство облегчает задачу взаимного согласования каскадов в многокаскадных усилителях Далее, при включении с общим эмиттером плоскостной транзистор позволяет получить большое усиление как по току, так и по напряжению.

На рис. 11-9 показана эквивалентная схема переменного тока для транзистора, включенного с общим эмит-

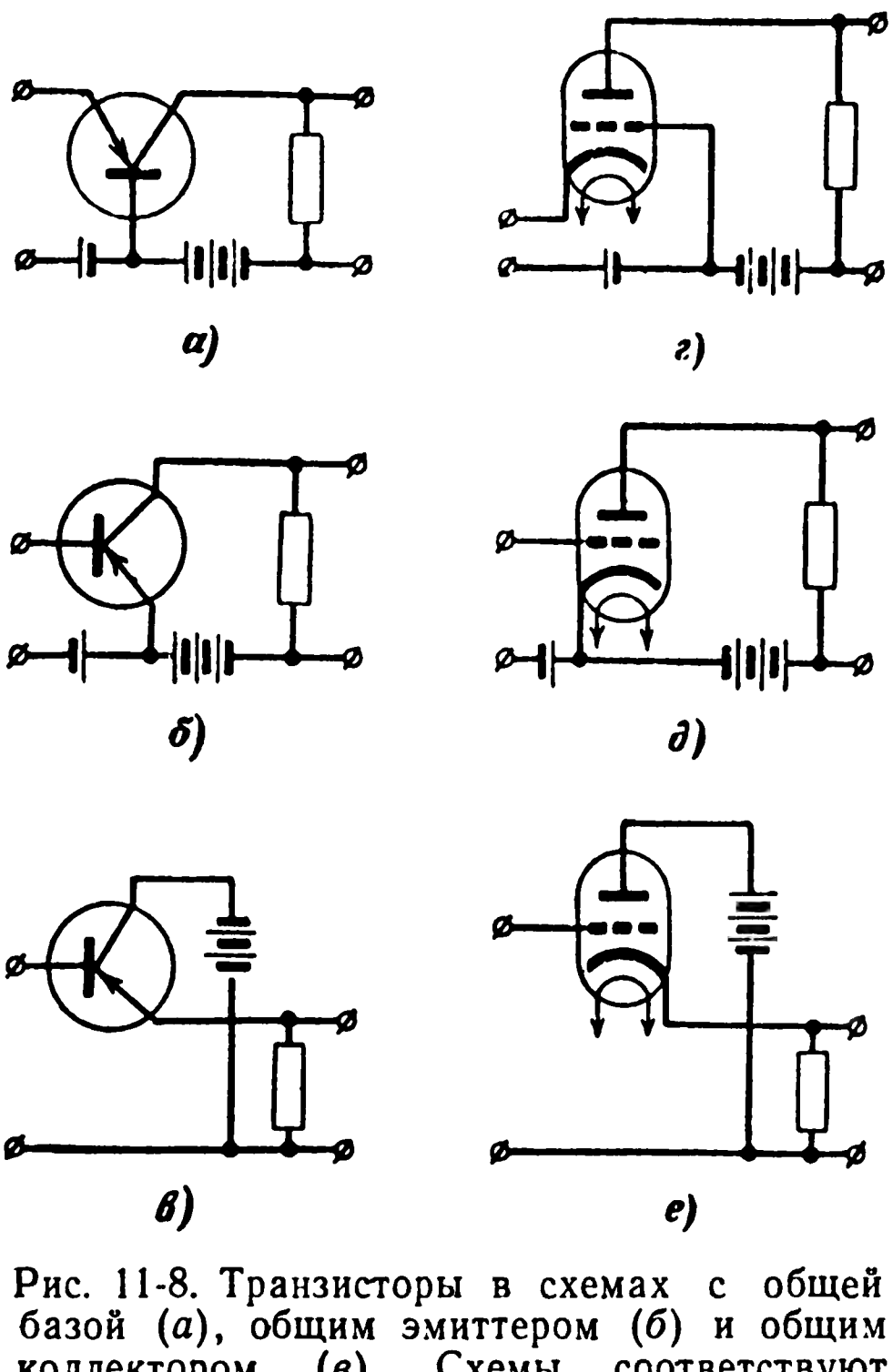


Рис. 11-8. Транзисторы в схемах с общей базой (*а*), общим эмиттером (*б*) и общим коллектором (*в*). Схемы соответствуют включению электронных ламп с общей сеткой (*г*), общим катодом (*д*) и общим анодом (катодный повторитель) (*е*).

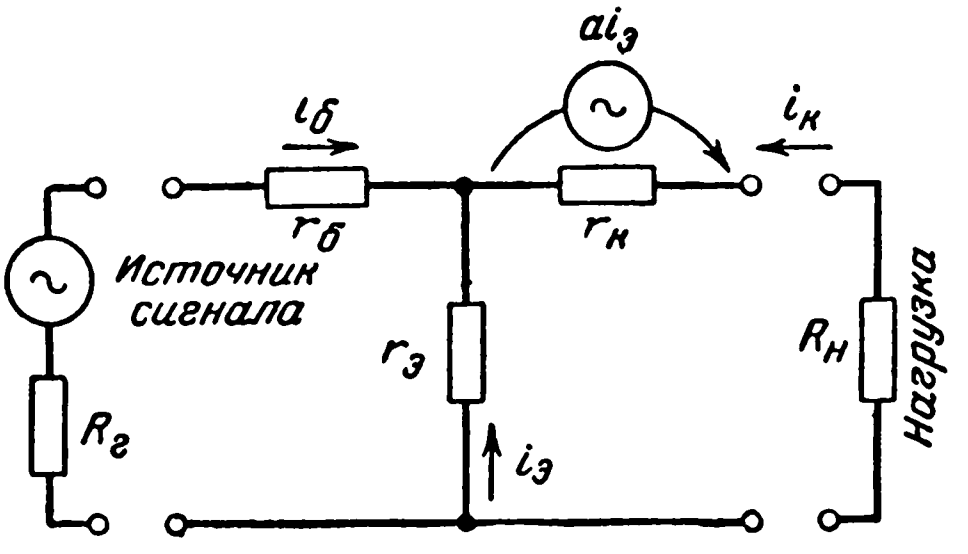


Рис. 11-9. Транзистор в схеме усилителя с общим эмиттером.

тером, работающего в качестве усилителя между генератором и нагрузкой. Входной ток теперь вводится как ток базы $i_б$ через базовое сопротивление эквивалентной Т-образной схемы. Анализ этой схемы показывает, что если даже внутренний коэффициент передачи тока a меньше единицы, отношение выходного тока к входному может быть во много раз больше. Так, анализ схемы дает:

Внешний коэффициент передачи тока $=$

$$=\frac{i_к}{i_б}=\frac{-r_э+ar_к}{r_э+(1-a)r_к+R_н}\approx\frac{ar_к}{(1-a)r_к+R_н}. \quad (11\text{-}17)$$

В частности, если $R_н$ уменьшается до нуля и коллектор, таким образом, замыкается по переменному току с общим эмиттером, то уравнение (11-16) сводится к следующему:

Внешний коэффициент передачи тока в режиме короткого замыкания $\approx$

$$\approx\frac{a}{1-a}\approx\frac{\alpha}{1-\alpha}. \quad (11\text{-}18)$$

Оно показывает, что при короткозамкнутом выходе возможно большое усиление по току. Если у транзистора $\alpha=0{,}99$, то его усиление по току в режиме короткого замыкания при включении с общим эмиттером составит 99.

Из уравнения (11-17) следует, что когда сопротивление нагрузки увеличивается, то усиление по току уменьшается. Однако при оптимальных согласованных с выходом нагрузках все же обычно достигается значительное усиление по току. Действительно, если каскады при включении транзисторов с общим эмиттером соединены без межкаскадных согласующих трансформаторов так, что выходная нагрузка одного каскада представляет собой входное полное сопротивление следующего каскада (самое большее — несколько сотен ом), то $R_н$ в уравнении (11-17) оказывается малым по сравнению с $(1—a)\ r_к$ и усиление по току на каскад практически равно усилению по току в режиме короткого замыкания, определяемому из уравнения (11-18).

Выводы, приведенные в этом параграфе, показывают, что чем α ближе к единице, тем большее усиление по току можно получить при включении транзистора по схеме с общим эмиттером. Это обстоятельство стимули-

ровало разработку транзисторов с большим α; последнего можно добиться уменьшением толщины базового слоя, чтобы сократить время прохождения неосновных носителей от эмиттера до коллектора и, следовательно, свести к минимуму рекомбинацию их по дороге. Кроме того, транзисторы для этих целей конструируют так, чтобы увеличить время жизни неосновных носителей в области базы.

Для транзисторов, включенных по схеме с общим эмиттером, соотношения между параметрами четырехполюсника и элементами эквивалентной схемы отличаются от соответствующих выражений для схемы с общей базой, которые были приведены на рис. 11-6.

Для включения с общим эмиттером, показанного на рис. 11-9, определение четырех частных производных при описании методом четырехполюсника приводит к следующей системе соотношений:

$$\left.\begin{aligned}
&\frac{\partial U_{\text{вх}}}{\partial I_{\text{вх}}} = z_{\text{вх}} = r_{\text{б}} + r_{\text{к}};\\
&\frac{\partial U_{\text{вх}}}{\partial I_{\text{вых}}} = z_{\text{о.прох}} = r_{\text{э}};\\
&\frac{\partial U_{\text{вых}}}{\partial I_{\text{вх}}} = z_{\text{п.прох}} = r_{\text{э}} - a r_{\text{к}};\\
&\frac{\partial U_{\text{вых}}}{\partial I_{\text{вых}}} = z_{\text{вых}} = r_{\text{э}} + (1-a)\, r_{\text{к}}.
\end{aligned}\right\} \qquad (11\text{-}19)$$

Далее, хотя графическое описание транзистора в том виде, в каком оно дано на рис. 11-2, вполне удовлетворительно, его форма неудобна для наиболее употребительных расчетов транзисторов в схемах с общим эмиттером. Характеристики транзисторов, предназначенные для анализа схем с общим эмиттером, обычно строятся в виде семейства коллекторно-эмитерных характеристик при различных значениях входных токов базы. Полнота описания может быть достигнута присоединением базо-эмитерных характеристик для различных токов коллектора. Однако в справочниках эти семейства базовых характеристик обычно опускаются аналогично тому, как в справочниках по электронным лампам часто приводятся только семейства анодных характеристик. На рис. 11-10 показано семейство коллекторных характеристик для транзистора с выращенным переходом, включенного по схеме с общим эмиттером. Заме-

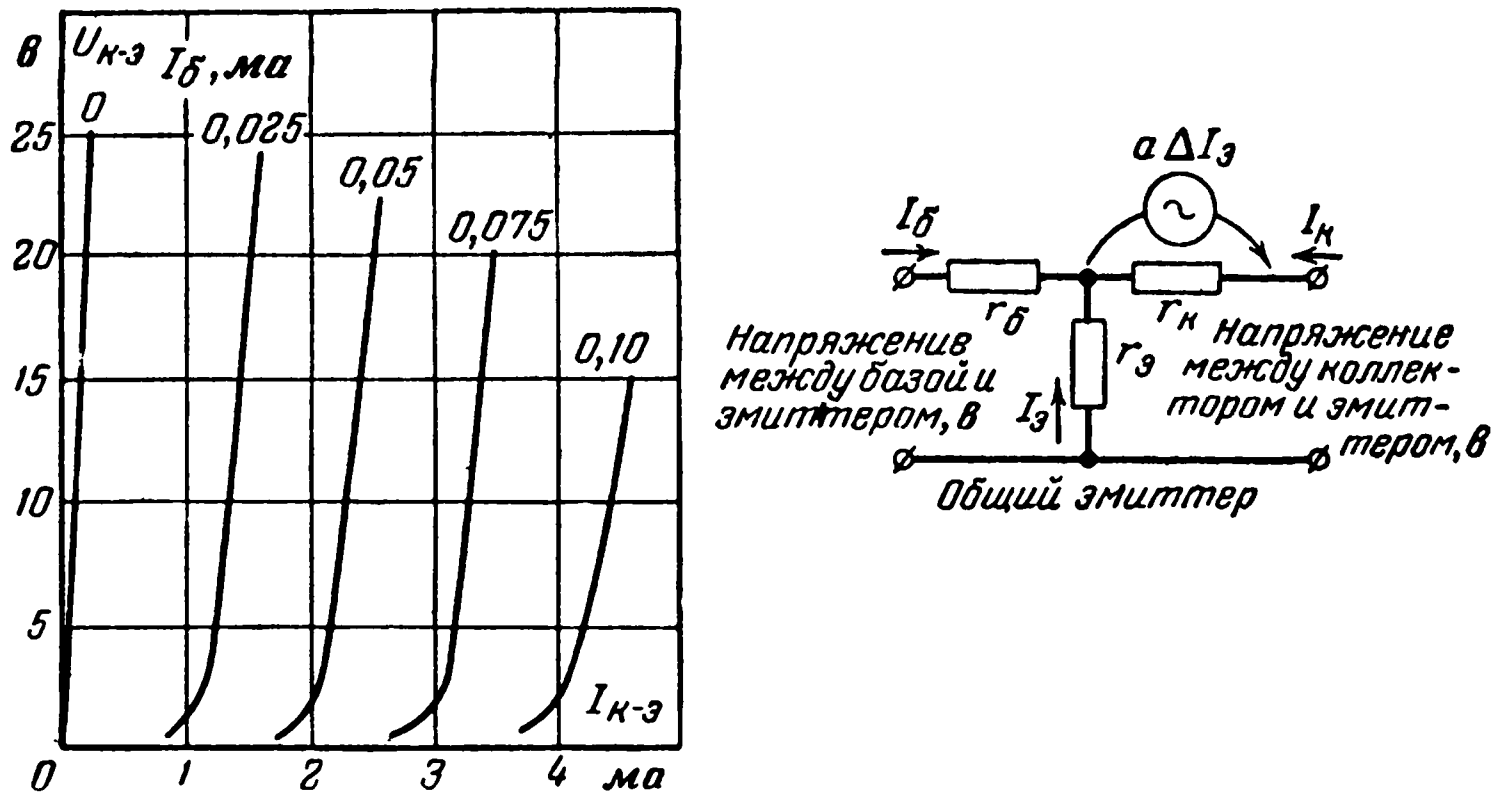


Рис 11-10 Выходные характеристики транзистора с *n-p-n* структурой при работе в схеме с общим эмиттером.

тим, что коэффициент передачи тока в режиме короткого замыкания по уравнению (11-18) можно определить из уравнения (11-10), найдя отношение

$$\left.\frac{\partial I_к}{\partial I_б}\right|_{U_к=\text{const}}$$

из соседних кривых графика для любой выбранной рабочей точки.

Для транзисторов, включенных по схеме с общим эмиттером, можно вывести следующие полезные выражения [Л. 4].

Входное полнсе сопротивление для выходной нагрузки $R_н$ равно:

$$r_б + r_э\left(\frac{r_к + R_н}{r_э + (1-a)\,r_к + R_н}\right). \qquad (11\text{-}20)$$

Выходное полное сопротивление для входного сопротивления генератора $R_г$ равно:

$$r_э + (1-a)\,r_к + \frac{r_э\,(ar_к - r_э)}{R_г + r_э + r_б}. \qquad (11\text{-}21)$$

Коэффициент усиления мощности при постоянных сопротивлениях генератора $R_г$ и нагрузки $R_н$

$$\frac{4R_г R_н\,(ar_к - r_э)^2}{\{(R_г + r_э + r_б)\,[r_э + (1-a)\,r_к + R_н] - r_э\,(r_э - ar_к)\}^2}. \qquad (11\text{-}22)$$

Максимальное усиление при оптимальных согласованных сопротивлениях генератора и нагрузки равно:

$$\frac{(r_э - ar_к^2)}{r_э r_к \left[\sqrt{\left(1+\frac{r_б}{r_э}\right)\left(\frac{r_э}{r_к}+1-a\right)}+\sqrt{1+\frac{r_б}{r_э}\left(\frac{r_э}{r_к}+1-a\right)}\right]^2}. \quad (11\text{-}23)$$

Включение транзистора по схеме с общим коллектором применяется в тех случаях, когда желательно работать при высоком сопротивлении генератора и малом сопротивлении нагрузки. В таких случаях транзистор подобен электронной лампе, работающей в схеме катодного повторителя, у которого входное полное сопротивление велико, а выходное — мало. Для транзистора при включении по схеме с общим коллектором соотношения между параметрами четырехполюсника и элементами эквивалентной схемы выражаются так:

$$\left.\begin{aligned}
&\frac{\partial U_{вх}}{\partial I_{вх}} = z_{вх} = r_б + r_к;\\
&\frac{\partial U_{вх}}{\partial I_{вых}} = z_{о.прох} = (1-a)\, r_к;\\
&\frac{\partial U_{вых}}{\partial I_{вх}} = z_{п.прох} = r_к;\\
&\frac{\partial U_{вых}}{\partial I_{вых}} = z_{вых} = r_э + (1-a)\, r_к.
\end{aligned}\right\} \quad (11\text{-}24)$$

Выражения для сопротивлений и коэффициента усиления для такой схемы включения имеют следующий характер.

Входное полное сопротивление при выходном сопротивлении нагрузки $R_н$ равно:

$$r_б + r_к - \frac{r_к^2 (1-a)}{r_э + (1-a)\, r_к + R_н}. \quad (11\text{-}25)$$

Выходное полное сопротивление при входном сопротивлении генератора $R_г$ равно:

$$r_э + (1-a)\, r_к - \frac{r_к^2 (1-a)}{R_г + r_б + r_к}. \quad (11\text{-}26)$$

Коэффициент усиления мощности при постоянных сопротивлениях генератора $R_г$ и нагрузки $R_н$ равен:

$$\frac{4R_г R_н r_к^2}{[(R_г + r_б + r_к)\{r_э + (1-a)r_к + R_н\} - r_к^2(1-a)]^2}. \qquad (11\text{-}27)$$

Максимальное усиление при оптимально согласованных сопротивлениях генератора и нагрузки равно:

$$\frac{\dfrac{r_к^2}{(r_б + r_к)\{r_э + (1-a)r_к\}}}{\left[1 + \sqrt{\dfrac{r_б}{r_б + r_к} + \dfrac{r_э r_к}{(r_б + r_к)\{r_э + (1-a)r_к\}}}\right]^2}. \qquad (11\text{-}28)$$

ЗАДАЧИ

1. Определите графически значения величин $z_{вх}$, $z_{о\,прох}$, $z_{п\,прох}$, $z_{вых}$ и α в рабочей точке, соответствующей напряжению коллектора — 25 *в* и току коллектора — 2 *ма* для транзистора, характеристики которого приведены на рис. 11-2. (Ответ: 600 *ом*; 250 *ом*; 56 000 *ом*; 30 000 *ом*; 1,9.)

2. Рассчитайте по этим величинам параметры $r_э$, $r_б$, $r_к$ и a эквивалентной схемы транзистора в этой же рабочей точке. Представляется ли величина $r_э$ «правдоподобной» в качестве сопротивления точечного выпрямителя при прямом токе 0,5 *ма*? «Правдоподобна» ли величина $r_к$ в качестве сопротивления точечного выпрямителя при обратном напряжении 25 *в*? Совпадает ли примерно значение a со значением α, определенным графически в задаче 1?

3. Определите графически коэффициент усиления мощности переменного тока для транзистора, включенного по схеме с общей базой и работающего в режиме класса *A* при малых сигналах и напряжении питания коллектора 50 *в*, сопротивлении нагрузки 12,5 *ком* и токе эмиттера 0,5 *ма*. Коэффициент усиления мощности определите как отношение мощности, выделяемой на нагрузке, к мощности, подводимой ко входу транзистора. (Ответ: около 85 или 19 *дб*.)

4 С помощью нагрузочной кривой, построенной в задаче 3, определите графически на эмитерных характеристиках входное полное сопротивление транзистора в указанной выше рабочей точке. Рассчитайте и сравните входное полное сопротивление в той же рабочей точке, используя параметры эквивалентной схемы, найденные в задаче 2.

5. Повторите задачу 1, используя характеристики, приведенные на рис. 11-3. Сравните результаты.

ЛИТЕРАТУРА

1. Peterson L. C, Equivalent Circuits of Linear Active Four-Terminal Networks, Bell System Technical Journal, 1948, vol. 27, p 593.

2 Ryder R M, Kircher R, Some Circuit Aspects of the Transistor, Bell System Technical Journal, 1949, vol 28, p 367

3 Wallace R L, Pietenpol W J, Some Circuit Properties and Applications of n—p—n Transistors, Bell System Technical Journal, 1951, vol 30, p 530

4 Ryder R M, Kircher R J, см [Л 2] этой главы

5 Лоу А и др Основы полупроводниковой электроники, Изд «Советское радио», 1958

6. Полупроводниковые приборы и их применение, сборник статей под редакцией Я А Федотова, вып 1—7, Изд «Советское радио» с 1956 г

ДОПОЛНИТЕЛЬНАЯ БИБЛИОГРАФИЯ

Читателям, которые заинтересуются некоторыми подробностями рассмотренных в этой главе вопросов, рекомендуем обратиться к следующим книгам для наведения справок или чтения параллельно с настоящей книгой

Bevitt, William D, Transistor Handbook, Prentice-Hall, Englewood Cliffs, N J, 1956

Кобленц А и Оуэнс Г, Транзисторы, теория и применение, Издательство Иностранной литературы, 1956

Lo A W, Endres R O, Transistor Electronics, Prentice-Hall, Englewood Cliffs, N J, 1956.

Martin, Thomas L, Electronics Circuits, Prentice-Hall, Englewood Cliffs, N J, 1955.

Ши Р, Полупроводниковые триоды и их применение, Госэнергоиздат, 1957.

ГЛАВА ДВЕНАДЦАТАЯ

РАБОТА ТРАНЗИСТОРОВ НА ВЫСОКИХ ЧАСТОТАХ И ВЫСОКОЧАСТОТНЫЕ ТРАНЗИСТОРЫ

В предыдущих двух главах мы рассмотрели два главных типа транзисторов с различными структурами, а именно: точечные и плоскостные, и сформулировали некоторые основные положения, касающиеся определения свойств транзисторов и их поведения в цепях. Теперь мы подготовлены к тому, чтобы перейти к рассмотрению наиболее существенного и определяющего параметра — частотной характеристики. Мы опишем причины зависимости работы транзисторов от частоты, обсудим физический механизм, ограничивающий полосу рабочих частот транзистора, а также проиллюстрируем примерами некоторые приемы, применяемые для расширения этой полосы при разработке высокочастотных транзисторов.

12-1. ЗАВИСИМОСТЬ α ОТ ЧАСТОТЫ

При работе транзистора в простой схеме усилителя (рис. 11-7 или 11-9), коэффициент усиления усилителя остается постоянным в пределах значительной полосы частот, начиная с нуля и кончая несколькими единицами или десятками килогерц, почти для всех типов транзисторов. Формулы для вычисления сопротивлений и коэффициентов усиления, которые мы привели в гл. 11, справедливы только в этой полосе частот. По мере повышения частоты коэффициент усиления падает, асимптотически приближаясь к линии с наклоном более 10 *дб* на декаду шкалы частот. Хотя начало этого спада может до известной степени изменяться путем надлежащего выбора входного и выходного полных сопротивлений, само наличие спада определяет верхнюю границу полосы частот, в пределах которой транзистор может работать в качестве усилителя с плоской частотной характеристикой, и, таким образом, ограничивает возможности использования транзисторов.

Одним из способов графического описания частотных ограничений транзистора является построение графика зависимости коэффициента передачи тока α в режиме короткого замыкания от частоты. Таким способом описывается поведение самого транзистора независимо от влияния внешней цепи на его частотную характеристику. На рис. 12-1 показаны кривая зависимости α от частоты для точечного транзистора, а также кривая, дающая величину сдвига фазы коэффициента α для той же полосы частот. Частота, при которой α падает до величины $\frac{1}{\sqrt{2}}$ ее значения при низкой частоте, называется верхней предельной частотой f_α. При этой частоте фазовый сдвиг $\alpha \approx 45°$.

Эти кривые могут быть вполне удовлетворительно представлены эмпирическим выражением

$$\alpha = \frac{\alpha_0}{1 + jf/f_\alpha}, \qquad (12\text{-}1)$$

где α_0 — значение α при низкой частоте;

f — частота;

f_α — верхняя предельная частота, которую мы только что определили.

Это выражение не учитывает сдвигов фазы более 90°, которые экспериментально наблюдаются при высоких частотах. Однако нас интересует в основном поведение транзистора в полосе частот ниже f_α, где уравнение (12-1) аналитически хорошо выражает реальные явления.

Физический механизм, обусловливающий это поведение, состоит в распозлании сгустков неосновных но-

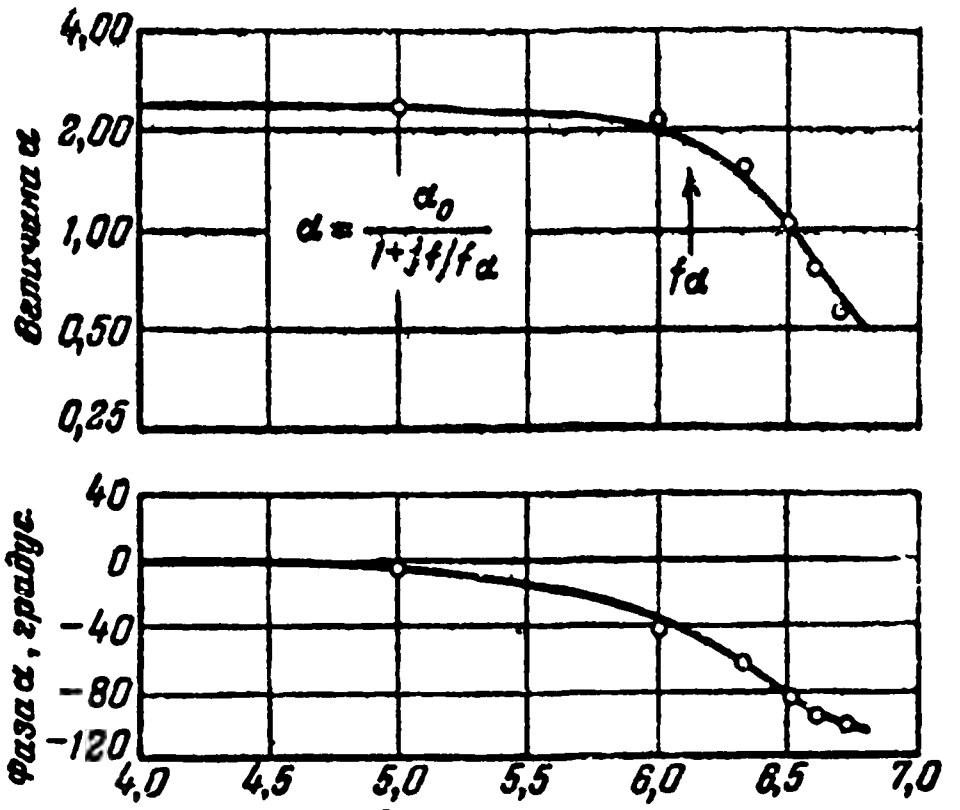


Рис 12-1 Зависимость величины и фазы коэффициента α точечного транзистора от частоты.

сителей зарядов в процессе их перемещения при диффузии внутри полупроводника от эмиттера к коллектору. Чтобы изучить этот механизм более детально, предположим, как это иллюстрируется рис. 12-2, что с эмитерного контакта в полупроводник выделилось небольшое облачко дырок в виде короткого прямоугольного импульса. Это облачко затем движется по направлению к коллектору частично за счет диффузии и частично за счет переноса электрическим полем коллектора, находящегося под обратным напряжением. По мере передвижения этого облачка оно расплывается. Некоторые дырки движутся по прямым путям и быстро достигают коллектора. Другие движутся по более искривленным путям и достигают коллектора с запозданием. Даже те частицы, коотрые движутся по почти

совпадающим путям, проходят весь путь в разные промежутки времени из-за случайных столкновений с другими частицами в процессе перемещения. Поэтому к тому времени, когда облачко достигает коллектора, оно оказывается сильно размазанным, его резкие края исчезают, что соответствует потере высокочастотных слагающих ряда Фурье для прямоугольного импульса. Если теперь предположить, что эмиттер испускает быструю последовательность таких облачков и если они

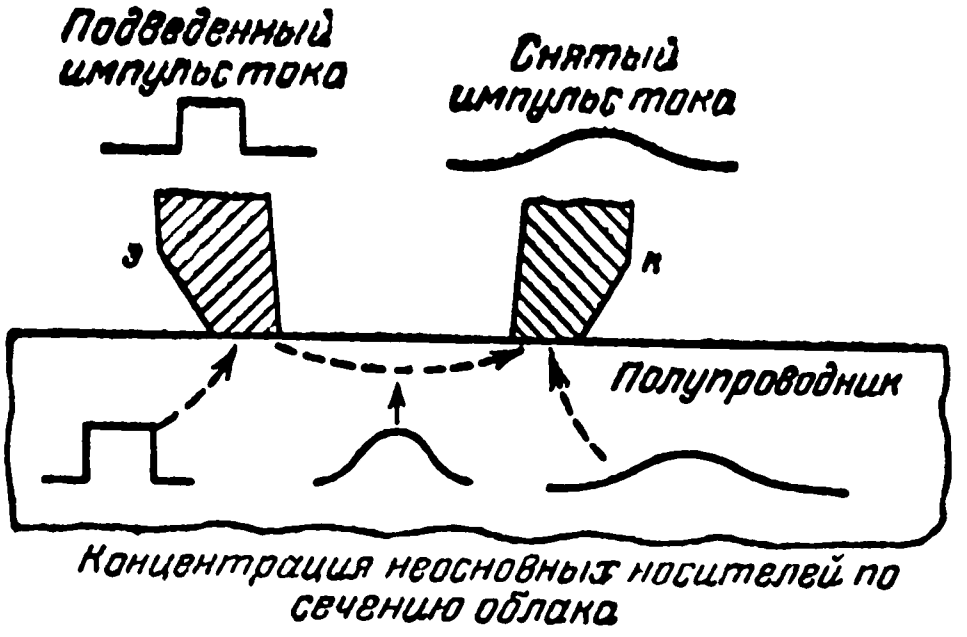


Рис. 12-2. По мере того как объемный заряд неосновных носителей, переносящий сигнал, перемещается от эмиттера к коллектору, он расплывается.

достаточно быстро следуют друг за другом, то к моменту их прихода к коллектору они в значительной степени перекроются. Результатом этого перекрытия является общее уменьшение амплитуды, тем большее, чем выше частота. Разумеется, это рассуждение применимо в такой же степени к синусоидальному сигналу на входе, как и к прямоугольному, который был нами взят лишь в качестве примера.

Сложная геометрия точечного транзистора не позволяет произвести простой теоретический расчет ожидаемого влияния разброса по времени прохождения. Однако в транзисторах с параллельно расположенными *p-n* переходами, в которых перемещение неосновных носителей через базовый слой происходит только за счет диффузии и в которых линии тока неосновных носителей на всех участках параллельны, такой расчет

можно выполнить, причем получается следующее выражение для верхней граничной частоты:

$$f_\alpha = \frac{D}{\pi W^2}, \qquad (12\text{-}2)$$

где D является коэффициентом диффузии[1] неосновных носителей в базовом слое и W — толщина базового слоя.

Если f_α выражено в мегагерцах и W — в микронах, то уравнение (12-2) приобретает вид:

для германиевого транзистора типа *n-p-n*

$$f_\alpha = \frac{4,9}{W^2}; \qquad (12\text{-}3)$$

для германиевого транзистора типа *p-n-p*

$$f_\alpha = \frac{2,3}{W^2}.$$

Эти формулы были тщательно проверены для транзисторов с выращенными переходами при максимально точном измерении толщины базового слоя. Проверка показала, что описанные выше физические процессы действительно являются причиной частотных ограничений в этих транзисторах.

Из уравнения (12-3) следует, что для создания транзистора, способного работать на высоких частотах, нужно сделать тонкий базовый слой. Кроме того, из него следует, что если делать базовый слой все тоньше и тоньше, то должна получиться неограниченно высокая предельная частота. Это заключение, действительно, правильно. Однако если постепенно делать базовый слой все более тонким, желая повысить значение f_α, то мы встретимся с другим эффектом, ограничивающим верхнюю рабочую частоту.

---

[1] Коэффициентом диффузии называется число частиц, диффундирующих за секунду сквозь поверхность площадью 1 *см*^{2}, расположенную нормально к градиенту концентрации частиц, имеющему значение 1 *частица*/*см*$^3 \cdot$ *см* В германиевом транзисторе типа *n-p-n* неосновными носителями заряда являются электроны, для которых $D = 94$ *см*^{2}/*сек* В германиевом транзисторе типа *p-n-p* неосновными носителями являются дырки, для которых $D = 44$ *см*^{2}/*сек*.

12-2. ВЛИЯНИЕ ЕМКОСТЕЙ ПЕРЕХОДОВ НА ЧАСТОТУ

Зависимость α от частоты является лишь одной из важных причин, ограничивающих верхнюю рабочую частоту транзистора. Другая причина связана с емкостями эмитерного и коллекторного переходов. Несмотря на то, что эквивалентные схемы транзисторов, показанные на рис. 11-5 и 11-9, удовлетворительны для анали-

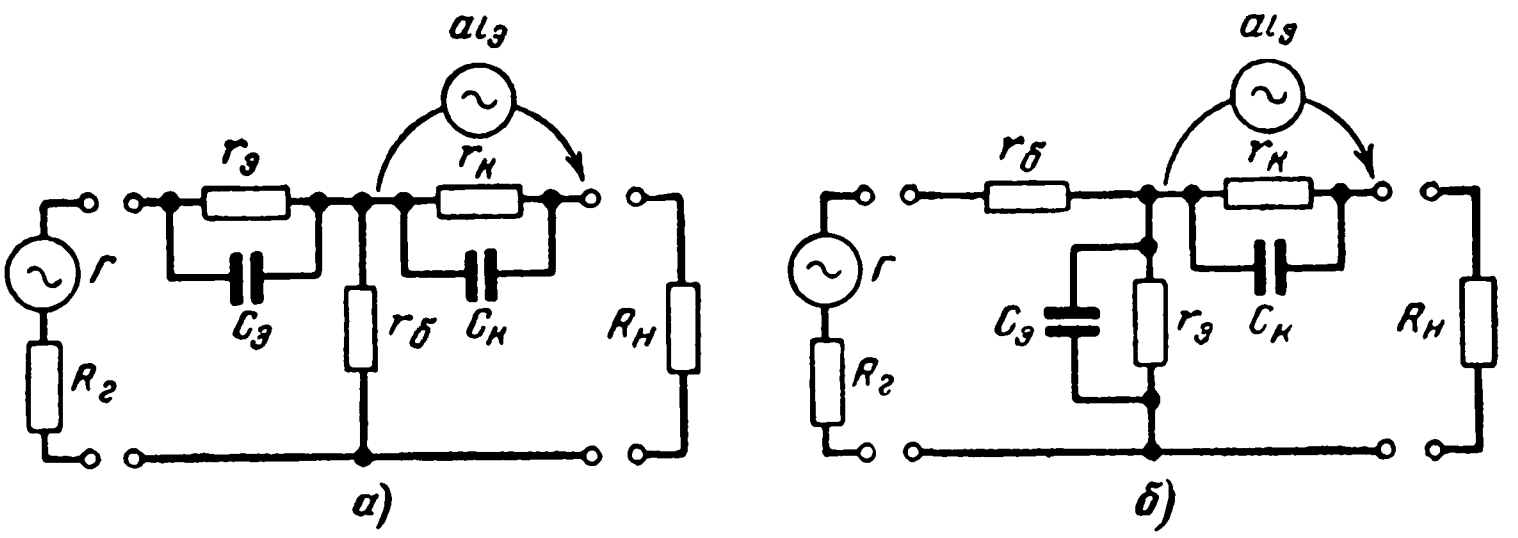


Рис. 12-3 Эквивалентная схема транзисторов для высокой частоты должна содержать емкости эмитерного и коллекторного переходов
а — общая база, *б* — общий эмиттер.

за поведения транзистора на низких частотах, указанные емкости переходов нужно принимать во внимание, если мы хотим составить эквивалентную схему, пригодную также и для анализа поведения транзистора на высоких частотах. Высокочастотные эквивалентные схемы транзисторов для случаев включения с общей базой и общим эмиттером приведены на рис. 12-3. В этих схемах сопротивления эмиттера и коллектора зашунтированы соответствующими емкостями переходов $C_э$ и $C_к$. Вспомнив выражения для сопротивлений и коэффициентов усиления, приведенные в предыдущей главе, можно сообразить, что введение этих емкостей в эквивалентные схемы транзисторов делает значения сопротивлений и коэффициентов усиления зависимыми от частоты.

Рассмотрим частотную зависимость максимально возможного коэффициента усиления мощности при согласованных входе и выходе, приведенного, например, в уравнении (11-6) для усилителя по схеме с общей базой.

Вместо $r_э$ в эти уравнения нужно подставить $r_э/(1+jr_эC_э\omega)$, выражающее сопротивление параллельно включенных $r_э$ и $C_э$. Подобным же образом мы

должны подставить вместо $r_к$ выражение $r_к/(1+jr_кC_к\omega)$. Выполнив это, а также произведя ряд значительных упрощений[1], мы получим для усилителя с общей базой:

Максимально возможный коэффициент усиления для плоскостного транзистора при включении с общей базой $\approx$

$$\approx \frac{\alpha^2 r_к}{4r_б} \cdot \frac{1}{1 + jr_кC_к\omega} . \qquad (12\text{-}4)$$

Нужно заметить, что в этом приближенном выражении $C_э$ не фигурирует. $C_э$ и $C_к$ имеют сравнимые значения, но $C_э$ шунтирует малое сопротивление $r_э$, тогда как $C_к$ шунтирует большое сопротивление $r_к$. Таким образом, влияние $C_э$ на пороговую частоту начнет сказываться только на частотах, много бóльших, чем те, при которых начинает сказываться $C_к$.

Мы видим, что это выражение для максимально возможного коэффициента усиления мощности обладает своей собственной зависимостью от частоты, совершенно отличной от зависимости α от нее. Зависимость максимально возможного коэффициента усиления мощности от частоты определяется последним множителем уравнения (12-4), который определяет предельную частоту f_G, соответствующую спаду усиления мощности на 3 *дб* относительно его максимального значения:

$$f_G \approx \frac{1}{2\pi r_кC_к} . \qquad (12\text{-}5)$$

Для типичного плоскостного тразистора, имеющего $r_к = 2$ *Мом* и $C_к = 10$ *пф*, предельная частота максимального коэффициента усиления мощности находится вблизи 8 *кгц*, несмотря на то, что верхняя предельная частота транзистора может равняться нескольким мегагерцам.

Зависимость этого коэффициента усиления мощности от частоты показана на рис. 12-4 сплошной линией. Коэффициент усиления достаточно постоянен в пределах полосы в несколько тысяч герц, спадая при часто-

[1] Более строгий расчет можно найти, например, в [Л. 7]. В приведенных выше выводах уравнения грубо аппроксимированы и вывод дан только для включения по схеме с общей базой. В нашем рассуждении мы пожертвовали строгостью расчета, чтобы упростить изложение основных идей.

тах, более высоких, чем предельная, с крутизной примерно в 10 *дб* на декаду. При частотах выше 1 *Мгц* начинает сказываться зависимость α от частоты и наклон кривой становится круче.

Если предполагается использовать транзистор при коэффициенте усиления мощности, меньшем максимального, то можно получить более высокую предель-

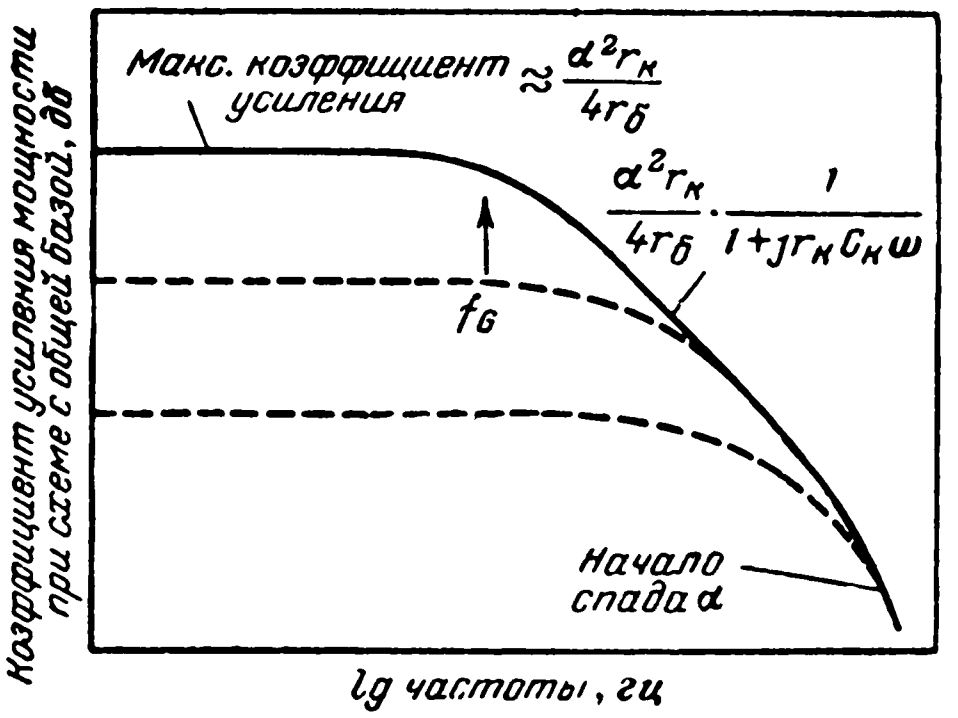


Рис 12-4 Зависимость усиления мощности от частоты в схеме усилителя с общей базой.

ную частоту. Такие случаи показаны пунктирными линиями на рис. 12-4. Таким образом, критерий широкополосности в полосе частот ниже верхней предельной частоты транзистора почти постоянен, вследствие чего можно расширить полосу частот за счет пропорционального уменьшения усиления. Подобные же выводы получаются для усилителей на транзисторах, включенных по схемам с общим эмиттером и общим коллектором.

Проанализируем уравнения (12-4) и (12-5) с точки зрения поиска путей для улучшения частотных характеристик усилителей на транзисторах. Критерий широкополосности для усилителя при включении с общей базой получается перемножением коэффициента усиления из уравнения (12-4) на предельную частоту из уравнения (12-5):

$$\textit{коэффициент широкополосности} \approx \frac{\alpha^2}{8\pi r_б C_к}. \qquad (12\text{-}6)$$

Из этого выражения видно, что коэффициент широкополосности может быть улучшен увеличением α или

уменьшением $r_б$ или $C_к$. Так как величина α в большинстве плоскостных транзисторов уже близка к единице, то путем увеличения α трудно намного улучшить коэффициент широкополосности. Напротив, уменьшение $r_б$ и $C_к$ вместе или порознь дает много возможностей для разработки транзисторов с повышенным коэффициентом широкополосности.

Из уравнения (12-6) также ясно, почему выигрыш, который можно получить за счет уменьшения толщины

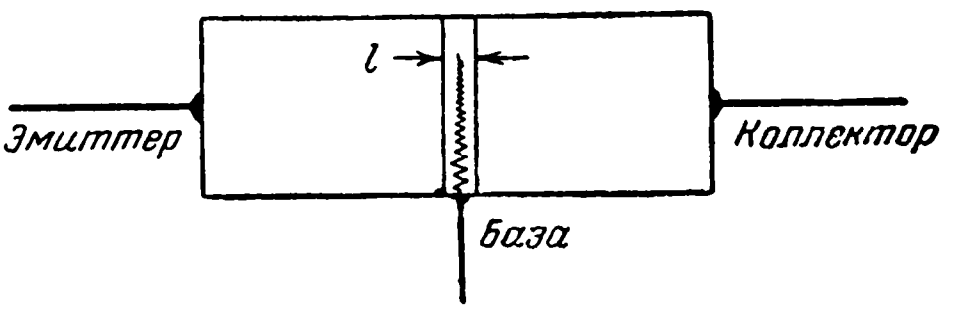


Рис 12-5 В эквивалентной схеме плоскостного транзистора некоторая часть сопротивления $r_б$ представляет собой сопротивление вдоль слоя от точки базового контакта до любой рассматриваемой точки слоя Уменьшение толщины слоя базы повышает эту составляющую сопротивления

базового слоя для повышения верхней предельной частоты, влечет за собой уменьшение коэффициента широкополосности. Это происходит потому, что бóльшая часть сопротивления, обозначенного в эквивалентной схеме транзистора через $r_б$, представляет собой распределенное сопротивление пленки базового слоя между точкой присоединения вывода базы и каждой точкой этого слоя, как показано на рис. 12-5. При уменьшении толщины базового слоя это сопротивление пропорционально увеличивается, так что хотя уменьшение толщины базового слоя действительно уменьшает влияние разброса времени передвижения и, таким образом, позволяет повысить f_α, оно также увеличивает $r_б$ и, значит, ухудшает коэффициент широкополосности транзистора. Можно попытаться устранить это ограничение, наращивая базовый слой из материалов с меньшим удельным сопротивлением. Однако это уменьшает напряжение обратного пробоя коллекторного перехода и, таким образом, ухудшает эксплуатационные качества транзистора. Кроме того, при этом понижается эффективность

ввода неосновных носителей из эмитерного перехода в базовый слой при протекании эмитерного тока.

Приведенные в последнем параграфе соображения обрисовывают некоторые из затруднений, возникающих при разработке транзисторов. Они также показывают, что разработчик постоянно стоит перед необходимостью компромисса между тем, что́ он хочет получить, и тем, чем он должен для этого пожертвовать.

Подводя итоги, мы можем заключить, что первоочередной задачей при разработке высокочастотного транзистора является обеспечение высокого коэффициента широкополосности путем уменьшения $r_б$ или $C_к$. Кроме того, поскольку диапазон частот, в пределах которого коэффициент широкополосности постоянен, ограничен на верхних частотах спадом α, возникает сопутствующая задача: повысить как можно больше верхнюю предельную частоту. Было проведено множество теоретических исследований и экспериментальных работ, чтобы преодолеть эти частотные ограничения или отдалить их влияние еще на несколько декад шкалы частот. В последующих параграфах изложено рациональное зерно этих изысканий и рассмотрены свойства приборов, построенных на этих принципах.

12-3. ПЛОСКОСТНЫЕ ТЕТРОДЫ

Одна из ранних успешных попыток обеспечить работу плоскостных транзисторов на высоких частотах привела к разработке тетрода [Л. 1]. В этом приборе базовый слой обычного плоскостного триода типа *n-p-n* снабжен вторым невыпрямляющим контактом на другой стороне полупроводникового образца. Транзистор включается в схему, как обычно, но на второй базовый контакт подается отрицательное смещение, как показана на рис. 12-6. При протекании постоянного тока вдоль базового слоя между контактами $б_1$ и $б_2$ этот ток создает падение напряжения IR, вследствие чего бо́льшая часть эмитерного перехода находится под обратным напряжением. Лишь небольшая часть эмитерного перехода, непосредственно примыкающая к базовому контакту $б_1$, находится под прямым напряжением и может эмитировать неосновные носители в базовый слой По-этому область прохождения тока между эмитером и

коллектором сужается до узкого канала возле контакта $б_1$, как показано на рис. 12-7.

В обычном плоскостном триоде типа *n-p-n* значительная часть сопротивления $r_б$ эквивалентной схемы

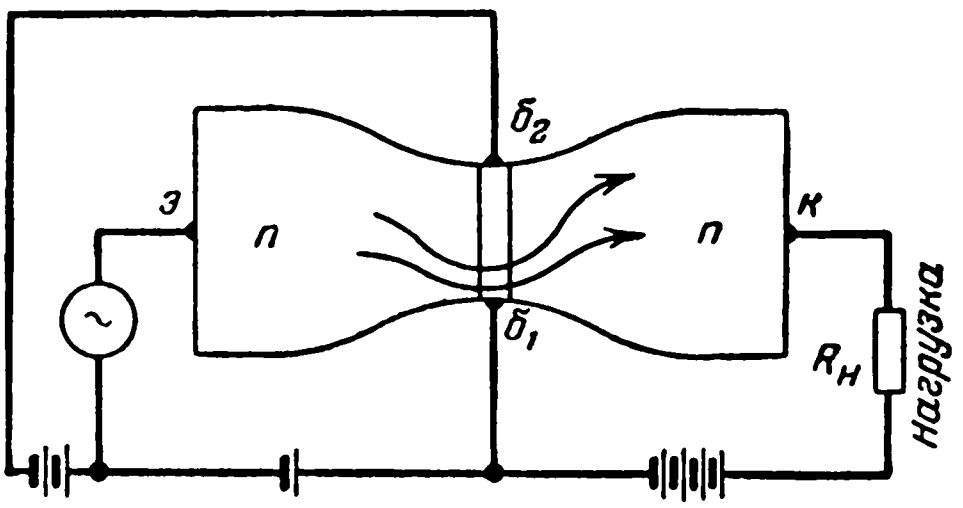


Рис 12-6 В тетроде с *n-p-n* структурой и двойной базой действие транзистора ограничено небольшой областью, примыкающей к главному контакту базы

представляет собой распределенное сопротивление базового слоя между точкой присоединения базового вывода и внутренней частью слоя. Однако в тетроде, поскольку рабочая часть базового слоя сужена до небольшого участка вблизи базового контакта, величина $r_б$ может быть на порядок меньше, чем у триода с такой же толщиной базового слоя. Чтобы использовать выго-

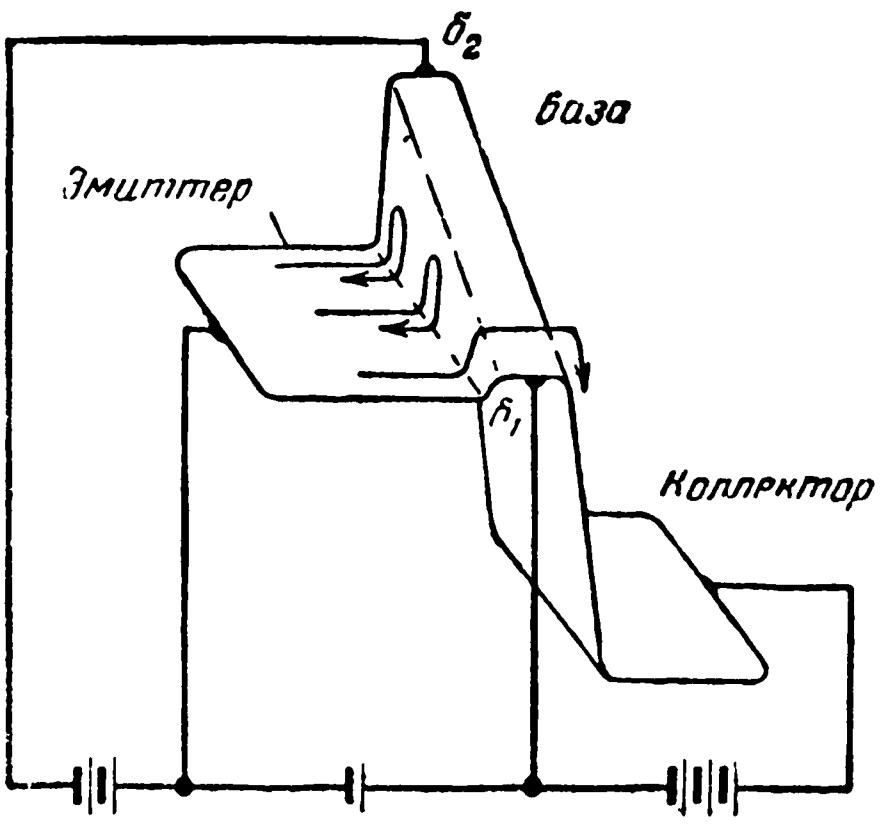


Рис. 12-7. Диаграмма энергий электронов в тетроде с *n-p-n* структурой Электроны могут пересекать эмитерный переход только вблизи контакта $б_1 б_2$.

ду повышения предельной частоты для α за счет уменьшения времени передвижения частиц, можно уменьшить толщину базового слоя тетрода до 0,005—0,0075 *мм* без опасения увеличить $r_б$ настолько, чтобы серьезно ухудшился коэффициент широкополосности прибора.

Первые изготовленные тетроды действительно могли работать на высоких частотах. В течение года были изготовлены тетроды, имевшие значение верхней предельной частоты для $\alpha = 15—20$ *Мгц* и способные генерировать колебания до 150 *Мгц*. В последующее время были введены некоторые усовершенствования, и в настоящее время возможно генерирование частот выше 1 000 *Мгц* [Л. 2].

Из-за небольшой величины рабочей площади тетрода ток коллектора у тетрода меньше, чем у триода с выращенным переходом, вследствие чего рабочая мощность тетродов всего 10—25 *мвт*.

12-4. ПОВЕРХНОСТНО-БАРЬЕРНЫЕ ТРИОДЫ

Значительным прогрессом в разработке новых типов триодов, имеющих весьма небольшое расстояние между эмиттером и коллектором и соответственно улучшен-

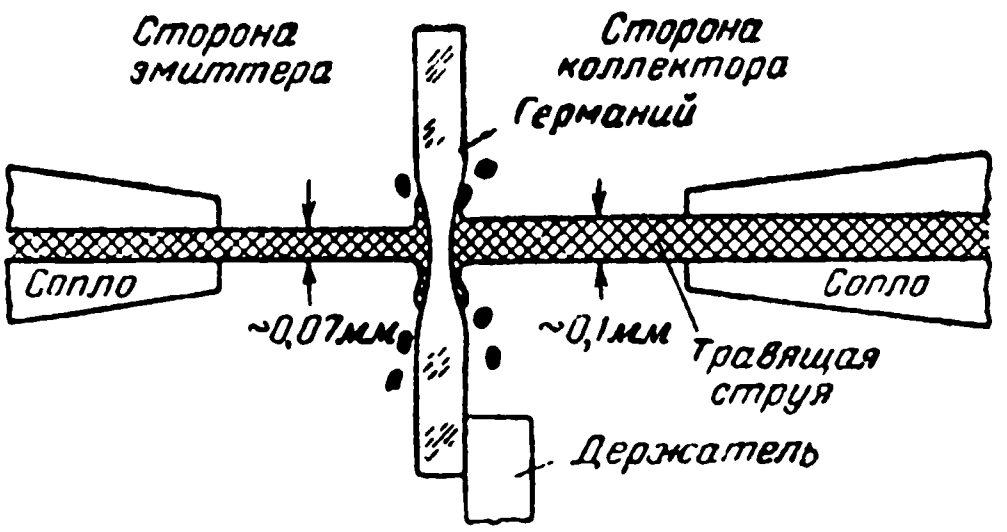


Рис 12-8 Контакты эмиттера и коллектора наносят на транзисторный поверхностно-барьерный элемент с помощью струй электролита

ные частотные характеристики, явилось создание поверхностно-барьерных транзисторов [Л. 3]. Для изготовления поверхностно-барьерного транзистора берется очень тонкая пластинка германия, толщина которой путем шлифовки и травления доводится до 0,075 *мм*. Затем эта пластинка устанавливается между двумя коаксиальными наконечниками (рис. 12-8), из которых

вытекают струи травящего раствора, вызывающего дальнейшее уменьшение толщины пластинки в центральном участке. Когда толщина оставшейся перемычки становится равной примерно 0,005 *мм*, на донышки обоих этих углублений наносятся электролитически контакты, причем теперь струи электролита используются как электроды. Пластины затем устанавливаются в держателе, осуществляющем контакт базы с пластинкой и имеющем эмитерный и коллекторный выводы, контактирующие с гальванически нанесенными электродами. Вся конструкция заключается в пластмассовую оболочку (рис. 12-9). Все устройство несколько напоминает конструкцию коаксиального точечного транзистора с утонченной в середине пластиной, показанную на рис. 10-5. Однако расстояния от эмиттера до коллектора в поверхностно-барьерном триоде в 10 раз меньше, чем в коаксиальном точечном.

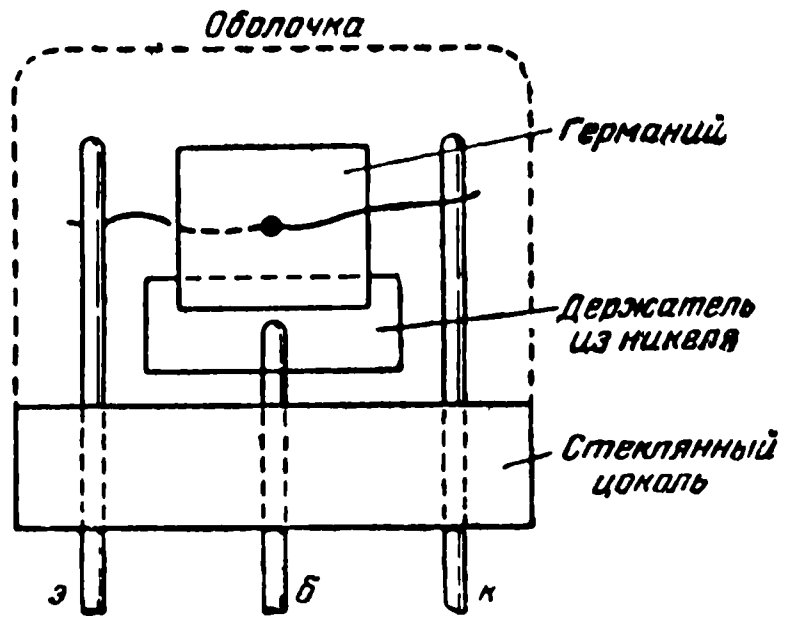


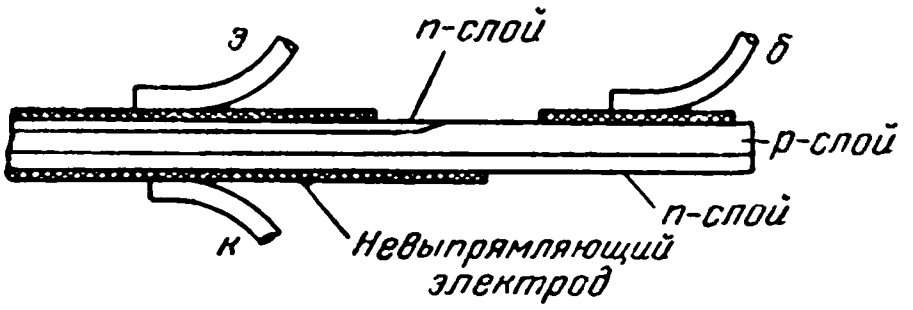
Рис. 12-9. Транзисторный поверхностно-барьерный элемент в герметической пластмассовой оболочке.

Поскольку в этих транзисторах базовый слой очень тонок, верхняя предельная частота их имеет величину от 30 до 50 *Мгц*. Несмотря на это, при такой конструкции тонкий базовый слой не приводит к недопустимо высокой величине $r_б$, как это было бы в обычном триоде с выращенным переходом, у которого вывод базы находится в контакте с базовым слоем только в одной точке. В поверхностно-барьерной конструкции контакт имеется по всему периметру вокруг участка германия, заключенного между эмитерным и коллекторным переходами, так что получается бесконечно большое число параллельных линий тока, идущих в центральный район базы. Таким образом, величина $r_б$ поддерживается в пределах нескольких сотен ом. Более того, поскольку рабочая часть прибора имеет столь ничтожные размеры, а площадь коллекторного контакта равна всего нескольким десятитысячным долям квадратного милли-

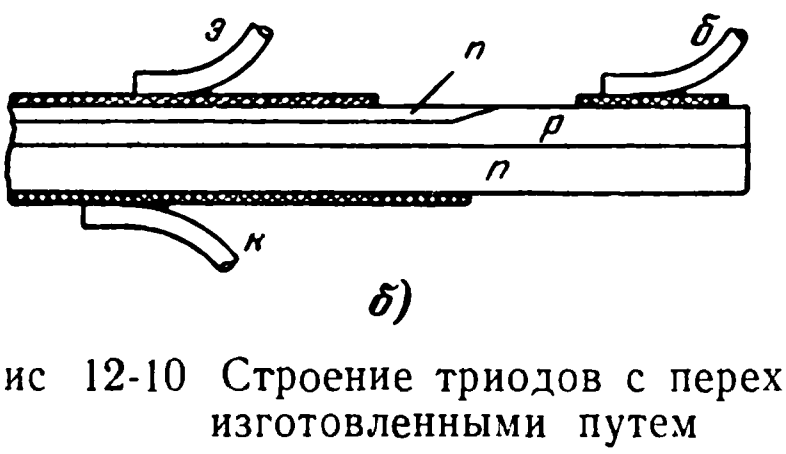
метра, величина $C_к$ не превышает 1—2 *пф*. Поэтому произведение $r_б C_к$ получается не выше, чем у обыкновенного выращенного триода, и в результате коэффициент широкополосности поверхностно-барьерного прибора получается повышенным, несмотря на очень малую толщину базового слоя.

12-5. ДИФФУЗИОННЫЕ ТРАНЗИСТОРЫ

Диффузионные методы изготовления *n-p* переходов вблизи поверхности полупроводника за счет обогащения этих поверхностей примесями описаны в гл. 8 в связи с изготовлением солнечных батарей. Диффузи-



а)



б)

Рис 12-10 Строение триодов с переходами, изготовленными путем

а — диффузии с противоположных сторон пластины типа *p*, *б* — поочередной (или даже одновременной) диффузии двух диффундирующих веществ в одну и ту же (верхнюю) сторону пластины типа *n*

онный метод делает возможным точное и контролируемое размещение переходов и, следовательно, осуществление контроля за расстояниями между переходами в приборах со многими переходами. На рис. 12-10,*а* показан возможный вариант диффузионного *n-p-n* транзистора. Такое строение переходов может быть осуществлено путем диффузии, например, фосфора с двух сторон пластинки *p*-типа до тех пор, пока оба перехода не подойдут весьма близко друг к другу. Успех этого метода определяется главным образом тем, насколько точно поверхности заготовки могут быть выполнены

плоскими и параллельными. Защитив надлежащим образом часть поверхности от насыщения ее примесью, можно обеспечить свободную поверхность исходного *p*-материала для создания с ним контакта. Прикрепить контакт к тонким диффузионным слоям трудно, но эта задача в настоящее время находится в стадии разрешения [1].

Другой возможный вариант диффузионного триода показан на рис. 12-10,*б*. Заготовкой является пластинка

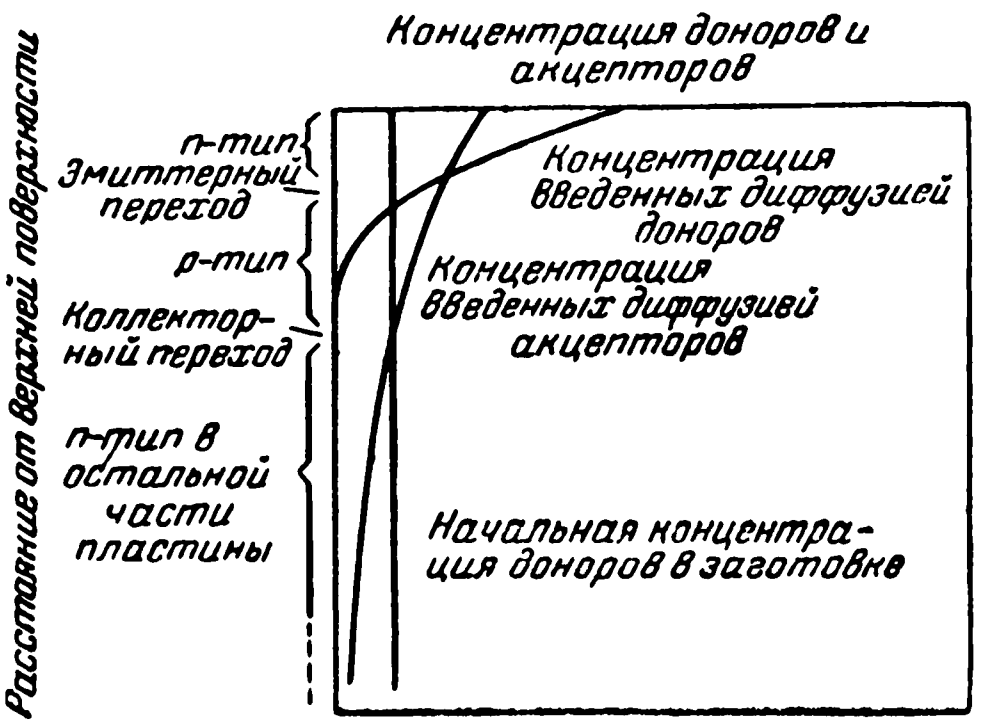


Рис. 12-11 В приборе, изготовленном путем двойной диффузии по рис 12-10,*б*, переход находится под поверхностью на глубине, где концентрации доноров и акцепторов равны

полупроводника *n*-типа, одна сторона которой насыщается примесью, вызывающей проводимость *p*-типа. Затем в ту же самую поверхность диффундирует примесь, вызывающая проводимость *n*-типа при более высокой поверхностной концентрации, но на меньшую глубину. Путем такой последовательной диффузии и при надлежащем контроле можно получить *p-n* переходы, отстоящие друг от друга только на сотые доли миллиметра и совершенно параллельные независимо от параллельности сторон заготовки. На рис. 12-11 показа-

---

[1] В настоящее время она разрешена, и диффузионные триоды стали одним из главных видов полупроводниковых триодов для высоких частот, вытеснив почти полностью тетроды, поверхностно-барьерные и точечные триоды и транзисторы с нейтральной прослойкой (*p-n-i-p* или *n-p-i-n*) *Прим ред*

ны кривые распределения концентраций примесей в таком приборе с двойной диффузией, из которых видно, как образуются переходы на тех глубинах, где концентрации донорных и акцепторных примесей оказываются одинаковыми.

Вариант конструктивного выполнения диффузионного плоскостного транзистора [Л. 4] показан на рис. 12-12. Пластина германия *p*-типа обрабатывается с одной стороны парáми мышьяка при повышенной

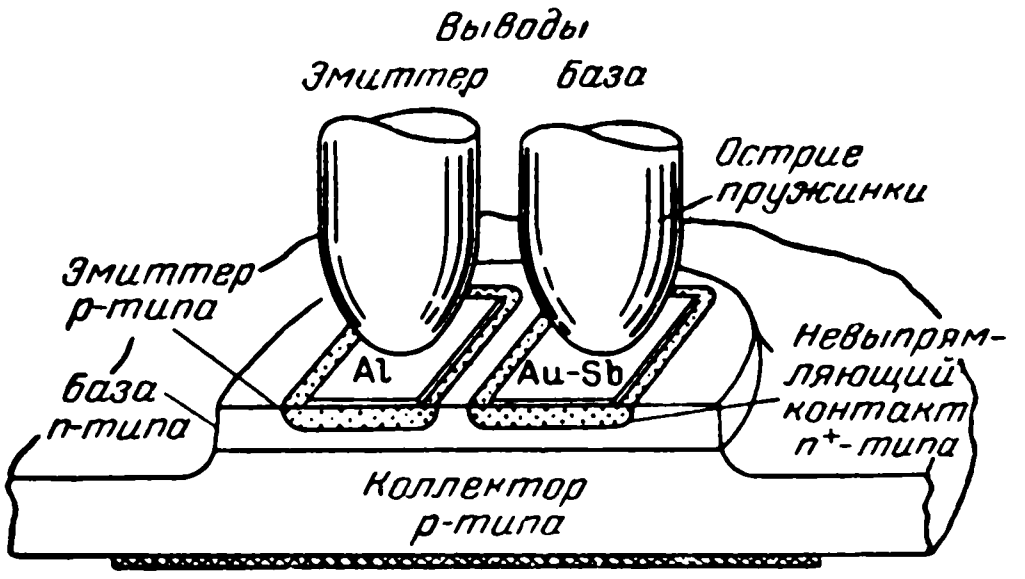


Рис 12-12 Разрез, показывающий строение *p-n-p* структуры с диффузионной базой Базовый слой, образованный путем диффузии, может иметь толщину порядка 1 *мк*.

температуре, пока на этой поверхности не образуется слой *n*-типа толщиной около 1 *мк*. *p-n* переход между этим слоем и массой заготовки является коллектором в готовом транзисторе. Затем посредством напыления в вакууме через соответствующие трафареты наносятся два металлических прямоугольных электрода, один из алюминия, а второй из сплава золота с сурьмой. После этого заготовка нагревается до такой температуры, при которой ионы металла диффундируют из этих напыленных пластинок в поверхностный *n*-слой, но не полностью сквозь него, а только на некоторую глубину. Таким образом, под алюминиевым электродом получается область эмиттера *p*-типа, а под электродом из сплава золота с сурьмой — область n^+-типа. Последняя служит только для обеспечения невыпрямляющего контакта с *n*-слоем, который становится базой транзистора. Затем прикрепляются контакты к телу заготовки и обоим металлическим электродам. Так получается

p-n-p структура, в которой эффективная толщина базового слоя между эмитерным и коллекторным переходами может быть сделана порядка 0,5 *мк*. Участок вокруг напыленных контактов закрывается капелькой лака, и вся остальная поверхность стравливается на глубину ниже уровня коллекторного перехода. В результате этой операции эмитерный и базовый электроды оказываются расположенными на небольшом возвышении, площадь которого определяет размеры коллектора и, таким образом, величину емкости коллектора. Все это устройство невероятно мало. Эмитерный и базовый электроды могут иметь размеры 0,025×0,05 *мм* и размещаются на расстоянии 0,025 *мм*. Коллекторное возвышение после травления может иметь диаметр всего 0,15 *мм*.

Ясно, что такой прибор является высокочастотным. Благодаря малой толщине базовой области верхние предельные частоты получаются около 500 *Мгц*. Кроме того, при такой конструкции сопротивление базы уменьшается ввиду образования кругового контакта по всей периферии вокруг рабочей области базы. Более того, так как коллекторное возвышение чрезвычайно мало, можно получить емкости коллекторного перехода, меньшие 1 *пф*. Поэтому с помощью такого транзистора была получена очень высокая верхняя предельная частота без ущерба для коэффициента широкополосности прибора.

На высокочастотные свойства этого транзистора с «диффузионной базой» благоприятно влияет еще одно обстоятельство, пока не рассмотренное нами и обусловленное природой самого процесса диффузии. Распределение примесей по толщине базового слоя неоднородно. Конструкция как доноров от первой диффузии, так и акцепторов от второй диффузии оказывается выше вблизи верхней поверхности возвышения. Эта неоднородность приводит к появлению постоянного внутреннего электрического поля по толщине базового слоя, облегчающего прохождение через этот слой неосновных носителей. Следовательно, время прохождения частиц через базу еще сокращается, уменьшается разброс этого времени и частотные ограничения отодвигаются за счет этого эффекта к более высоким частотам. Теория этого внутреннего электрического поля рассмотрена

в гл. 19. Приборы, в которых такое «вмороженное» электрическое поле используется для облегчения продвижения частиц, называют «дрейфовыми транзисторами» [Л. 5]. Насколько именно этот эффект улучшает высокочастотные характеристики транзистора с диффузионной базой, в точности оценить нельзя, поскольку нужно принимать во внимание еще множество других факторов, определяющих высокочастотные свойства прибора.

12-6. ТРАНЗИСТОРЫ С КОЛЛЕКТОРНОЙ НЕЙТРАЛЬНОЙ ПРОСЛОЙКОЙ *p-n-i-p* ИЛИ *n-p-i-n*

Существует конструкция транзисторов, в которой поддерживается низкая величина произведения $r_б C_к$ при одновременном использовании преимуществ очень тонкого базового слоя. Это достигается у транзисторов с нейтральной прослойкой *n-p-i-n* или *p-n-i-p* [Л. 6]. В таких конструкциях произведение $r_б C_к$ уменьшается главным образом за счет уменьшения $C_к$.

Предположим, что в область коллекторного перехода введена пленка диэлектрического материала так, чтобы увеличить эффективную толщину перехода. При этом его емкость уменьшается. Если можно было бы изготовить плоскостной транзистор со следующим чередованием типов проводимостей: *n-p*-собственная-*n*, то он имел бы как раз такое строение. На рис. 12-13 показаны схемы транзисторов типов *n-p-i-n* и *n-p-n*, причем тонкий коллекторный переходный слой последнего показан растянутым так, чтобы он включил в себя всю область с собственной проводимостью первого.

Современная технология изготовления транзисторов типа *n-p-i-p* или транзисторов с собственным переходом включает сочетание диффузионных и сплавных методов. Изготовляются по большей части приборы типа *p-n-i-p*, требующие более простых технологических приемов, чем *n-p-i-n*. На рис. 12-14 показан схематический разрез выпускаемого промышленностью транзистора типа *p-n-i-p*. Для изготовления прибора служит тонкая заготовка из германия с собственной проводимостью. Первой операцией является придание одной из поверхностей проводимости *n*-типа путем насыщения ее мышьяком на глубину около 0,025 *мм*. Этот диффузионный слой в дальнейшем является базой. В него с поверхно-

сти (не насквозь) вводится эмитерный слой *p*-типа путем вплавления индия. Подобным же образом в противоположную сторону заготовки вплавляется коллекторный слой *p*-типа. К базовому слою приплавляется кольцевой невыпрямляющий контакт.

В структуре типа *p-n-i-p* достигается толщина базовой области 0,005 *мм* между сплавным эмитерным слоем типа *p* и исходным материалом с собственной проводимостью. Верхняя предельная частота достигает порядка 100 *Мгц*. Поскольку толщина коллекторного перехода велика, так как она включает в себя всю толщину собственного слоя, $C_к$ уменьшается примерно до 0,5 *пф*. Кроме того, выполнение базового контакта в виде кольца, а также возможность изготовления базового слоя *n*-типа с малым удельным сопротивлением позволяют уменьшить $r_б$ примерно втрое по сравнению с величиной, обычно получающейся в транзисторах с выра-

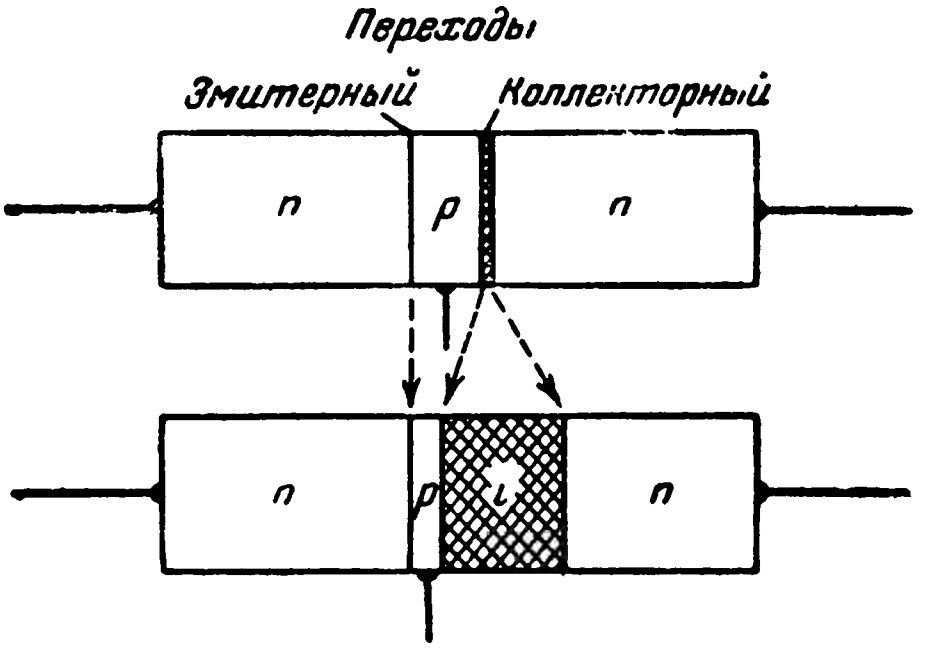


Рис. 12-13. В *n-p-i-n* структуре коллекторный запирающий слой занимает всю толщину слоя с собственной проводимостью

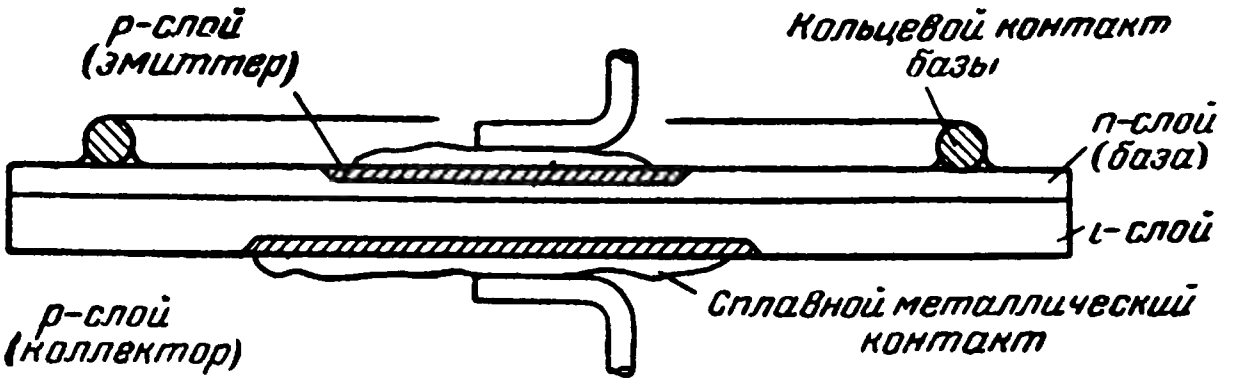


Рис. 12-14 Транзистор с коллекторной нейтральной прослойкой *p-n-i-p*

щенными переходами. Таким образом, коэффициент широкополосности усилителя, ограничиваемый произведением $r_6 C_к$ путем использования структуры *p-n-i-p* может быть улучшен примерно на два порядка по величине.

ЛИТЕРАТУРА

1. Wallace R. L., Schimpf L. G., Dickten E., A Junction Transistor Tetrode for High-Frequency Use, Proceedings of the Institute of Radio Engineers, 1952. vol. 40, p. 1395.
2. Wallace R. L., Junction Tetrode Transistor, Bell Laboratories Record, 1955, vol. 33, p. 121.
3. The Surface-Barrier Transistor, Proceedings of the Institute of Radio Engineers, 1953, vol. 41, p. 1702
4. Lee C. A., A High-Frequency Diffused-Base Germanium Transistor, Bell System Technical Journal, 1956, vol. 35, p. 23.
5. Kromer H., Zur Theorie des Diffusions und des Drifttransistors, Archiv der elektrischen Übertragung, 1954, vol. 8, p. 223.
6. Early J. M., pnip and npin Junction Transistor Triodes, Bell System Technical Journal, 1954, vol. 33, p. 517.
7. Wallace R. L., Pietenpol W. J., Some Circuit Properties and Applications of n-p-n Transistors, Bell System Technical Journal, 1951, vol. 30, p. 530.
8. Полупроводниковые приборы и их применение, сборники статей под редакцией Я. А. Федотова, вып. 1—7, Изд. «Советское радио», с 1956 г.

ГЛАВА ТРИНАДЦАТАЯ

ПРОЧИЕ ТИПЫ ТРАНЗИСТОРОВ

В последних трех главах мы рассмотрели конструкции, свойства и методы описания серийно выпускаемых транзисторов, а также транзисторов, разработанных для работы на высоких частотах. В нашем обзоре не рассмотрено еще несколько различных типов транзисторов, описанию которых и посвящена настоящая глава.

13-1. СТЕРЖНЕВОЙ ТРАНЗИСТОР С МОДУЛИРОВАННОЙ ПРОВОДИМОСТЬЮ

В гл. 5 описана идея использования ввода неосновных носителей для уменьшения удельного сопротивления полупроводника и получения этим путем повышенной проводимости диода в прямом направлении. Исто-

рически впервые идея модуляции проводимости была реализована в стержневом транзисторе [Л. 1], схематически показанном на рис. 13-1. Из германия *n*-типа сделана заготовка, имеющая узкий мостик в средней части, как показано на схеме, сечением примерно 0,125×0,125 *мм* и длиной 1,25 *мм*. Концы заготовки расширены, чтобы облегчить присоединение невыпрямляющих выводных контактов. Такая заготовка получается следующим образом: тонкая германиевая пластинка приклеивается к стеклянному основанию, сред-

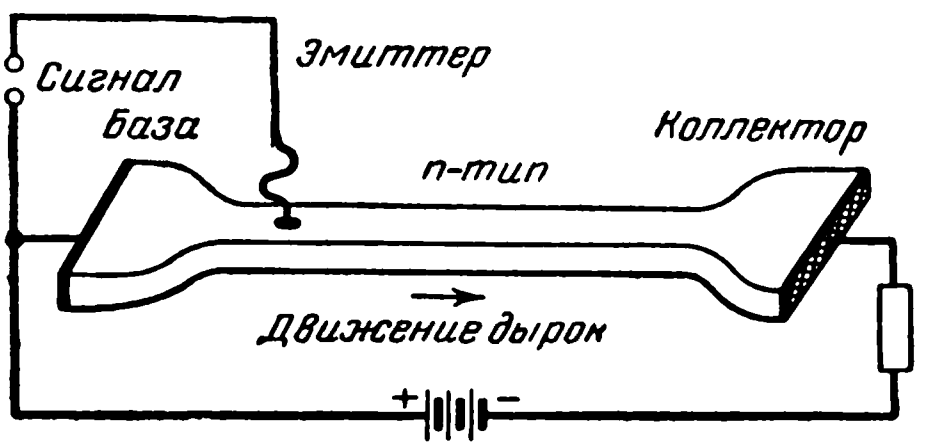


Рис 13-1 Стержневой триод с модулированной проводимостью, представляющий собой полупроводниковый мостик между проводящими контактами с эмиттером у одного конца

няя часть пластинки закрывается шаблоном нужной формы, после чего лишний материал удаляется пескоструйным аппаратом. Вблизи базового конца стержневой части полупроводника размещается точечный эмитерный контакт.

При разомкнутой цепи эмиттера электроды движутся от коллектора к базе через стержень высокого сопротивления. Если цепь эмиттера замкнута, то у эмитерного контакта в стержне выделяются положительные дырки, которые дрейфуют к коллектору, уменьшая сопротивление стержня, в результате чего в цепи между коллектором и базой может протекать большой ток. К тому же на каждую дырку, вводимую в стержень, из концевых выводов должен быть введен один добавочный электрон, чтобы нейтрализовать пространственный заряд дырок. Эти добавочные электроны также играют свою роль в увеличении проводимости стержня.

Такой прибор может усиливать мощность, так как

высокое сопротивление стержня модулируется током частиц, инъектируемых эмиттером, имеющим малое сопротивление. Этот прибор может работать на частотах. которые лишь немногим выше частот звукового диапазона. Сравнительно большое расстояние дрейфа между эмиттером и коллектором, длительное время передвижения приводят к низкому значению верхней предельной частоты. Уменьшение размеров прибора для улучшения его работы на высоких частотах имеет физические пределы. Ограничена также и величина отсасы-

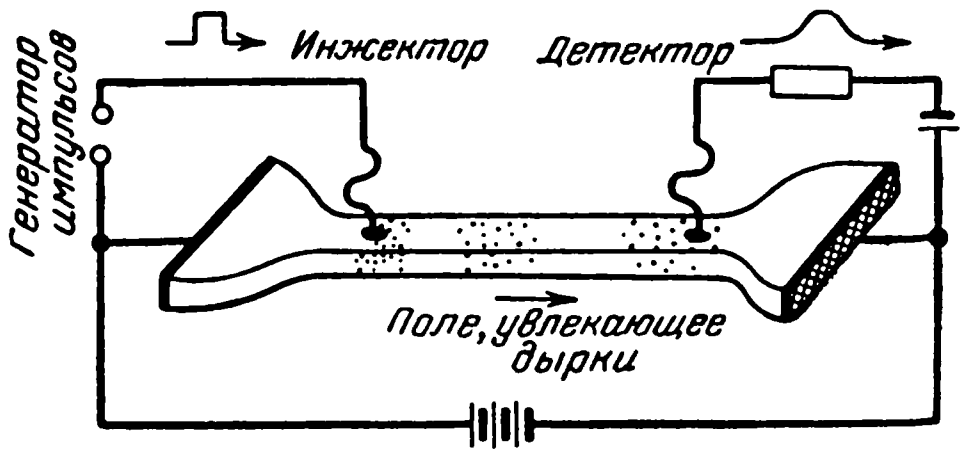


Рис 13-2 Полупроводниковый стержень в качестве линии задержки. Точками показан объемный заряд дырок, дрейфующих вправо.

вающего поля, которое можно создать в стержне для уменьшения времени прохождения частиц в связи с опасностью пережечь стержень. Поэтому возможности стержневых усилителей ограничены и разработке этих приборов уделялось мало внимания.

Были исследованы возможности таких конструкций в качестве линий задержки для использования в элементах памяти вычислительных устройств. Схема такой «линии задержки» показана на рис. 13-2. Сгустки неосновных носителей, образующиеся в стержне при импульсах тока в цепи эмиттера, дрейфуют вдоль стержня и улавливаются коллекторным контактом, расположенным на некотором расстоянии. На рисунке показаны три таких скопления дырок, дрейфующих вдоль стержня и расплывающихся по дороге. К сожалению, анализ показывает, что в стержне одновременно может существовать не более трех или четырех таких скоплений неосновных носителей без опасности их взаимного перекрытия за счет расплывания; в противном случае им-

пульсы на коллекторе становятся неразличимыми. Кроме того, время сохранения импульса ограничивается временем нейтрализации заряда неосновных носителей за счет рекомбинации (обычно порядка десятых долей микросекунды).

13-2. КАНАЛЬНЫЕ ТРАНЗИСТОРЫ

Выше уже упоминалось, что выпрямительный барьер становится толще, когда к нему прикладывается обратное напряжение. Это явление утолщения барьера было использовано для управления током в приборах,

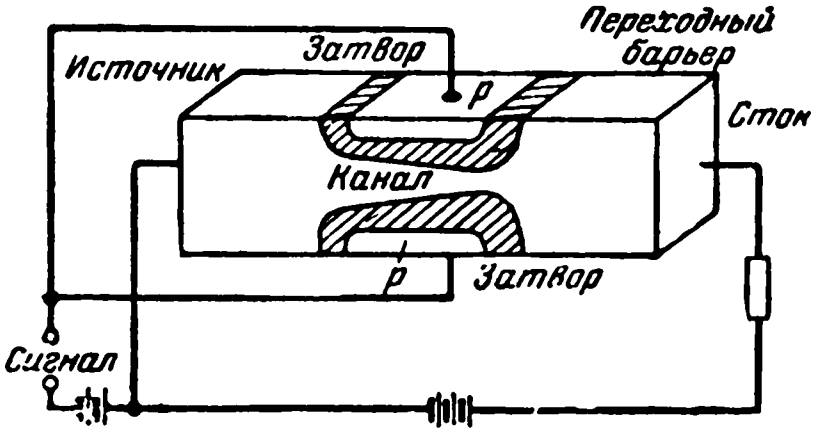


Рис 13-3 В канальном транзисторе ток проходит через канал, ширина которого определяется мгновенным напряжением на затворе его

названных «филдистерами» [Л. 2 и 3] или канальными транзисторами [Л. 4 и 5]. На рис. 13-3 показана схема канального транзистора. Он изготовляется из стержня полупроводника *n*-типа наплавлением участков *p*-типа. Весь прибор невелик по сравнению с обычным транзистором. Заготовка может иметь толщину около 0,125 *мм*, а наплавленные участки *p*-типа — протяженность около 0,125 *мм* вдоль стержня и вплавлены в него на глубину 0,05 *мм*. Левый конец стержня называется источником, правый — стоком, а два вплавленных участка *p*-типа — затвором. На схеме показаны также электрические цепи при включении прибора в качестве усилителя.

Канальный транзистор работает следующим образом. Если входной сигнал отсутствует, электроны входят в германий у источника, проходят через канал внутри затвора и выходят к стоку. На затворы подается постоянное обратное смещение, так что переходы здесь «выпячиваются» в канал, частично его перекры-

вая. Изменение величины сигнала, поданного на затвор, вызывает изменение толщины этих барьеров в области затвора. Таким образом, ширина канала, через который должен протекать ток от источника к стоку, также меняется при изменении входного сигнала, в результате чего происходит модуляция тока. Требуемая мощность входного сигнала переменного тока невелика и должна быть достаточной лишь для перезарядки емкостей переходов через сопротивление канала. Рассеиваемая в затворе мощность постоянного тока также невелика, так как через барьеры затвора протекает лишь обратный ток утечки. На рис. 13-3 переходы показаны более толстыми около конца, обращенного в сторону стока, чем с другой стороны. Это происходит потому, что напряжение между затвором и каналом на этом конце выше на величину падения напряжения на длине канала.

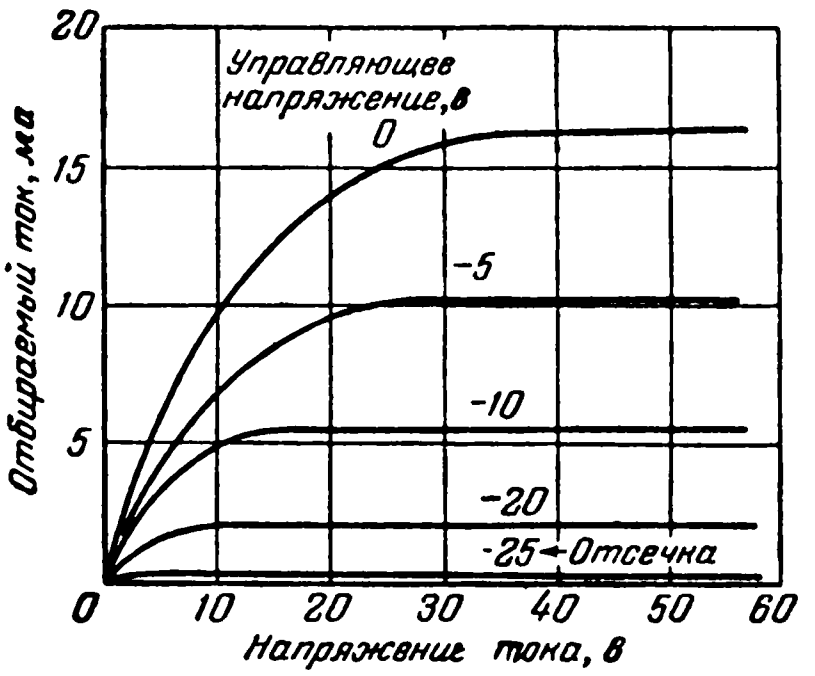


Рис 13-4 Семейство характеристик канального транзистора

Следует особо подчеркнуть, что при работе канального транзистора не происходит перемещения неосновных носителей; поэтому имеются разногласия в вопросе о том, можно ли назвать этот прибор настоящим транзистором. Поскольку он работает с помощью только одного вида заряженных носителей, его часто называют униполярным прибором. Из всех рассмотренных нами твердых усилительных приборов канальный транзистор представляет собой наиболее близкую аналогию с электронным триодом. Затвор действует как управляющая сетка в том отношении, что напряжение, приложенное к затвору, определяет ширину канала, через который проходит ток от источника к стоку, точно так же как напряжение на управляющей сетке определяет ширину потенциальной седловины, через которую должен проходить ток между катодом и анодом электронной лампы.

На рис. 13-4 показано семейство характеристик при-

бора как функции напряжения стока для различных значений напряжения затвора. Как и можно было ожидать, при достаточном отрицательном напряжении затвора можно заставить барьеры затвора расшириться настолько, чтобы перекрыть канал почти полностью, что приводит к запиранию тока. Можно также отметить, что при достаточно высоком напряжении стока все кривые на рис. 13-4 приближаются к состоянию насыщения. Это происходит потому, что как только напряжение стока превысит определенную величину, барьеры затвора начинают утолщаться в сторону стока, а не в сторону канала. Вследствие этого канал удлиняется, а не сужается и повышенное напряжение стока проявляется в виде повышенного падения напряжения на увеличившейся длине канала.

В настоящее время изготовляются канальные транзисторы с током насыщения стока от 10 до 20 *ма* при нулевом смещении затвора и крутизной характеристики 1—2 *ма/в* в области участка насыщения. Имеется, однако, явный источник искажений, проявляющийся в неодинаковом расстоянии между кривыми семейства этих характеристик. С помощью этого транзистора удавалось генерировать колебания частотой около 50 *Мгц*.

13-3. ЛАВИННЫЕ ТРАНЗИСТОРЫ

Из гл. 5 мы уже знаем, что если обратное напряжение, приложенное к *p-n* переходу, увеличивается выше пределов, в которых наблюдается эффект насыщения обратного тока, то происходит пробой перехода, проявляющийся в возникновении резкого увеличения тока при небольшом увеличении обратного напряжения. Это явление наблюдается также и в характеристиках плоскостных транзисторов. В обычном режиме транзистора напряжение коллектора поддерживается ниже уровня, при котором возникает пробой. Однако можно наблюдать несколько необычное поведение цепи, если работать при напряжении коллектора ниже уровня пробоя, но в той области, где происходит лавинное умножение тока коллектора. На рис. 13-5 показано семейство характеристик сплавного транзистора типа *p-n-p* в области увеличенных напряжений коллектора, включая область лавинного умножения. У транзисторов, предназначенных для работы в этой области, напряже-

ние пробоя должно быть ниже, чем у обычных транзисторов. Прибор, показанный на рис. 13-5, изготовлен с учетом этого требования. Его напряжение пробоя равно примерно 28 *в*. Существенно отметить, что при высоких напряжениях коллектора в области, непосредственно предшествующей пробою, $\Delta I_к$ превышает $\Delta I_э$. В этом диапазоне напряжений коллектора прибор усиливает ток цепи эмиттер — коллектор как при постоянном напряжении, так и при переменных сигналах.

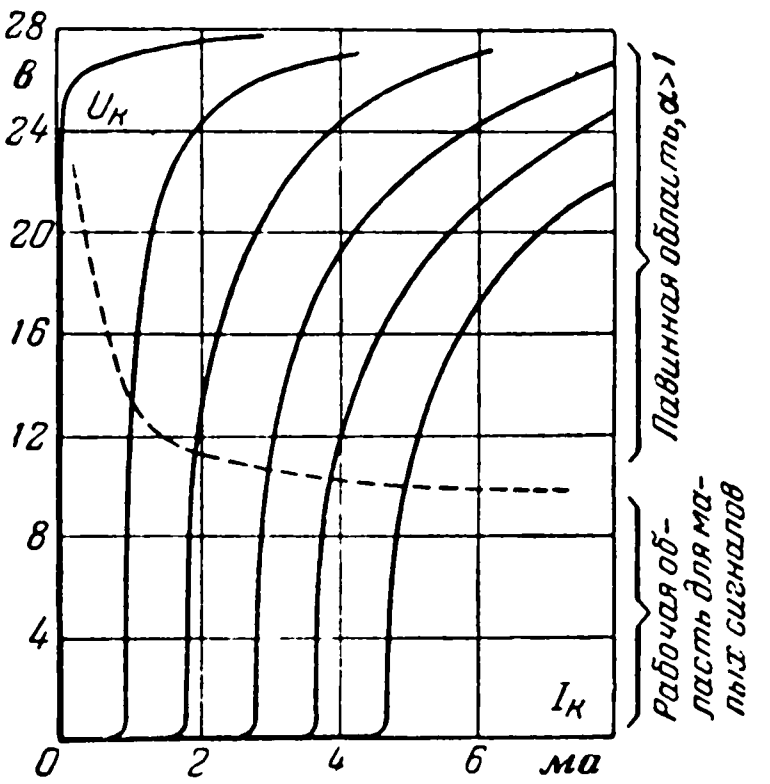


Рис 13-5 Увеличение коэффициента α свыше единицы для лавинного транзистора при достаточном увеличении напряжения коллектора.

Если этот прибор включить в цепь, изображенную на рис. 13-6,*а*, то в характеристике цепи обнаружатся некоторые особенности, показанные на рис. 13-6,*б* и напоминающие характеристики газоразрядного прибора. Это явление можно объяснить с помощью рис. 13-6,*а*, на котором направление положительного тока обозначено стрелками.

Между коллектором и базой постоянно протекает небольшой ток, обозначенный $I_{к0}$. Этот ток представляет собой обычный диодный обратный ток утечки коллекторного перехода при обратном напряжении. Протекая через внешнее сопротивление базы, этот ток создает небольшое прямое смещение на эмитерном переходе. Поэтому с эмиттера в базовый слой протекает ток $I_э$. Часть этого эмитированного тока $\alpha I_э$ течет через коллекторный зажим, тогда как другая $(1—\alpha)\ I_э$ проходит через вывод базы, имея направление, противоположное $I_{к0}$, и, таким образом, уменьшая прямое сопротивление эмиттера. Эмитерный ток поддерживается при этом небольшим до тех пор, пока α меньше единицы. Однако если напряжение коллектора увеличить настолько, чтобы α стало *больше* единицы, то направление тока $(1—\alpha)I_э$ меняется, *увеличивая* прямое смеще-

ние эмиттера и повышая эмитерный ток за счет эффекта положительной обратной связи. При этом величина выходного тока ограничивается только общим омическим сопротивлением цепи и напряжением коллектора, необходимого для того, чтобы поддерживать значение α при этом токе немного бóльшим единицы. На рис. 13-6,*б* этот процесс соответствует смещению вверх нагрузочной линии, первоначально пересекавшей в точке *А* вос-

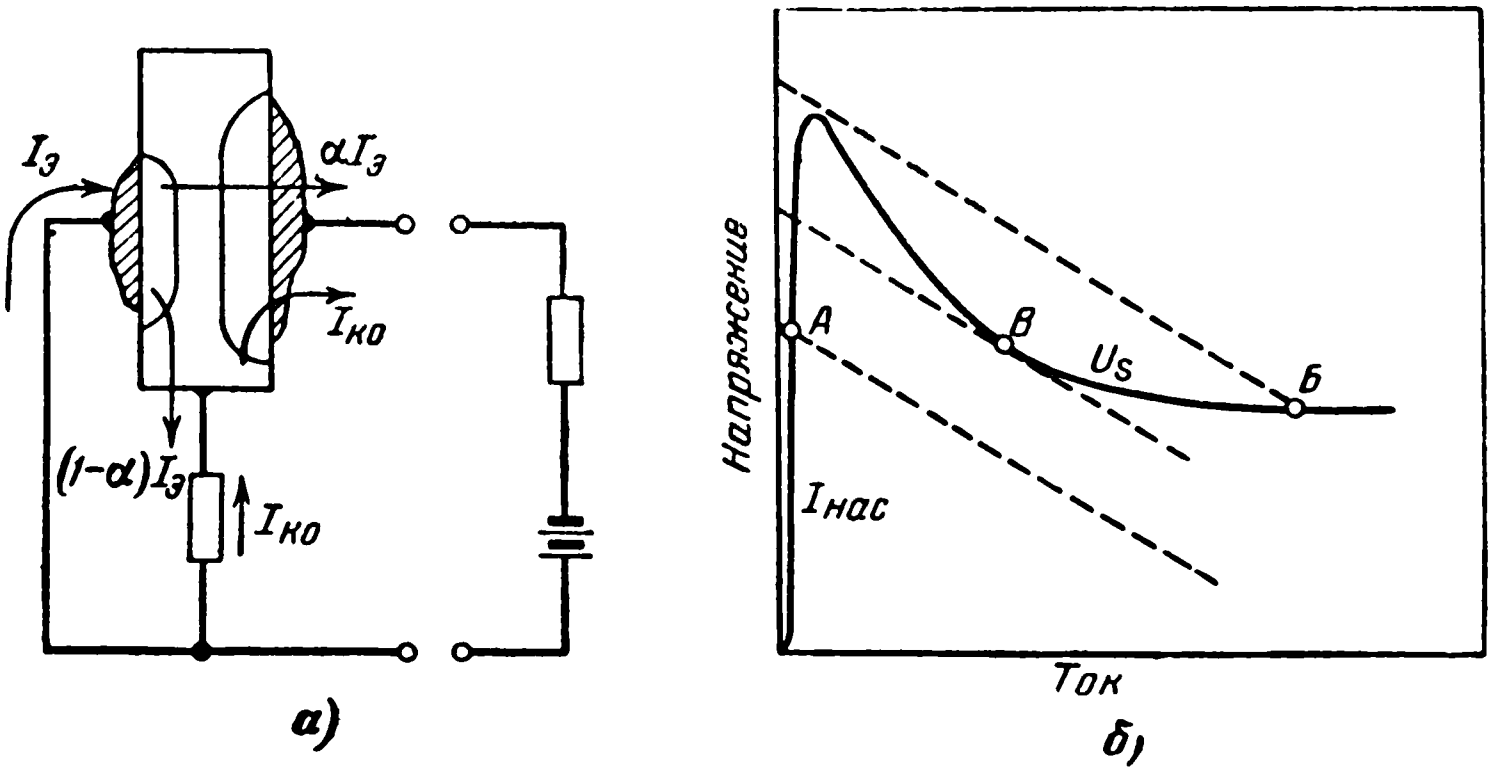


Рис 13-6 Лавинный транзистор (*а*) и его характеристики (*б*).

ходящую часть характеристики, пока рабочая точка не пройдет через максимум кривой и стабилизируется при режиме большого тока в точке *Б*. Транзистор может быть возвращен в первоначальный режим с высоким сопротивлением только путем снятия на короткое время коллекторного напряжения или понижения его настолько, чтобы сместить нагрузочную линию ниже точки *b* характеристики.

Характеристика, показанная на рис. 13-6,*б*, имеет некоторые особенности, обеспечивающие преимущество цепи с лавинным транзистором перед цепью с газоразрядным прибором. Напряжение пробоя может быть выбрано по желанию в пределах от 10 до 50 *в*. Установившееся напряжение не превышает нескольких вольт, в результате чего получается значительная разность между напряжением пробоя и напряжением установившегося режима. Для серийного плоскостного транзистора в такой цепи при включении получается ток до 100 *ма*. Физические явления при усилении тока в ла-

винном транзисторе подобны имеющим место при электрическом разряде в газе [Л. 7]. Бóльшая часть напряжения, приложенного к коллектору, падает на коллекторном переходе. Этот переход очень тонок, и внутри него образуется весьма сильное электрическое поле (10^5 *в/см*). Электрон, попавший в этот переход, больше приобретает энергии от поля, чем теряет ее за счет возбуждения тепловых колебаний кристаллической решетки. Таким образом, прежде чем он дойдет до другой стороны перехода, он получит достаточную избыточную энергию, чтобы ионизовать один из атомов решетки. При этом освобождаются новый свободный электрон и новая дырка. Если поле достаточно сильно, чтобы сделать α бóльшим единицы, то возникает лавинный пробой между эмиттером и коллектором.

13-4. ЛОВУШЕЧНЫЕ ТРИОДЫ И ТРИОДЫ ТИПА *p-n-p-n*

При разборе различных теорий, с помощью которых можно объяснить тот факт, что в точечных транзисторах $\alpha > 1$, было высказано мнение [Л. 8], что если изготовить четырехслойный транзистор, подобный изображенному на рис. 13-7, то он также должен иметь $\alpha > 1$.

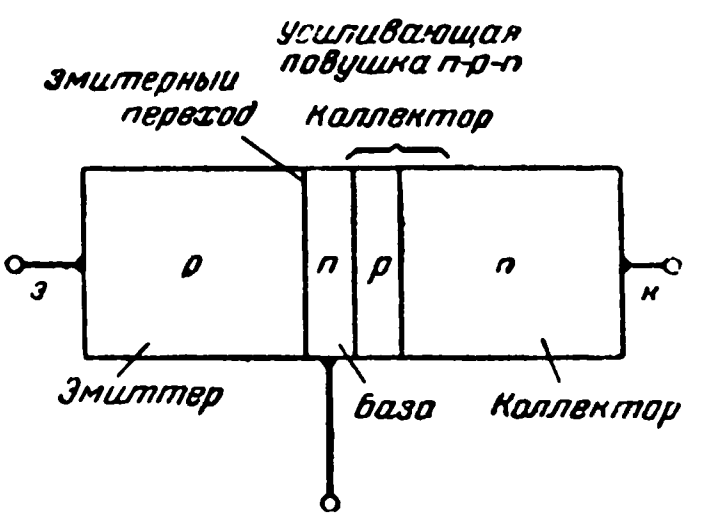


Рис 13-7. Структура перехода, при которой $\alpha > 1$

С помощью технологических приемов того времени такой прибор нельзя было изготовить, однако по мере развития технологии изготовления выращенных и сплавных переходов стало возможным изготовление триодов типов *n-p-n* и *p-n-p* и были исследованы конструкции транзисторов [Л. 9], структурно эквивалентные изображенным на рис. 13-7. Один из них показан на рис. 13-8. Он является просто составным прибором, полученным путем соединения двух транзисторов типов *p-n-p* и *n-p-n* так, как показано на схеме. Можно показать, что такое устройство имеет эффективное значение

$$\alpha \approx \frac{\alpha_{p\text{-}n\text{-}p}}{1 - \alpha_{n\text{-}p\text{-}n}}.$$

Любой прибор, имеющий в рабочем режиме коэффициент передачи тока больше единицы, может быть использован как ключ [Л. 11]. Поэтому и описанный выше прибор может быть также применен для этих целей.

С развитием диффузионных методов изготовления переходов оказалось практически целесообразным выполнять приборы типа *p-n-p-n* [Л. 11] в виде одного куска германия или кремния, применяя сочетание методов диффузии и сплавления для создания нескольких имеющихся в них переходов. На рис. 13-9,*а* показана симметричная структура типа *p-n-p-n*, центральный переход которой изготовлен методом диффузии, а остальные переходы—путем вплавления в поверхность заготовки. Этот прибор может работать в цепи, подобной показанной на рисунке, как двухполюсный ключ без внешнего вывода базы.

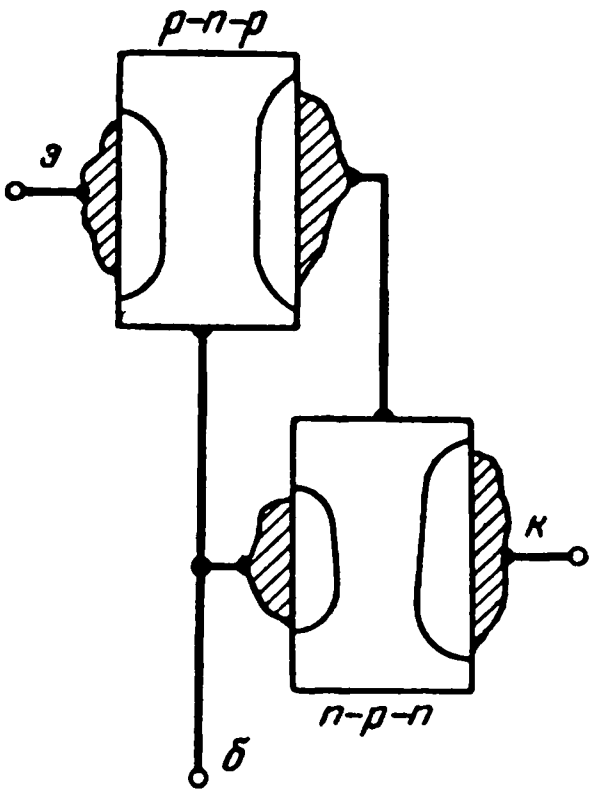


Рис. 13-8. Комбинация сплавных транзисторов типов *n-p-n* и *p-n-p*, эквивалентная элементу, изображенному на рис 13-7.

В подобной ключевой цепи полярность смещения выбирается так, что при низкой величине напряжения прибор работает во многом подобно простой цепи из трех последовательно соединенных плоскостных выпрямителей со средним переходом, находящимся под обратным напряжением, и двумя крайними — под прямым напряжением. В таком режиме ток, пропускаемый прибором, практически является обратным током насыщения утечки среднего перехода, соответствующим крутой части характеристики на рис. 13-9,*б*. Этот ток настолько мал, что соответствующее прямое напряжение на крайних переходах имеет величину, меньшую, чем обычно используемая для плоскостных транзисторов, работающих в качестве усилителей класса А.

Прибор типа *p-n-p-n* можно рассматривать как комбинацию двух отдельных транзисторов, обозначен-

ных на рис. 13-9 № 1 и 2, с перекрывающимися областями структур и свободными потенциалами баз. Иначе его можно рассматривать как соединение двух транзисторов с двумя отдельными эмиттерами, работающими на общий коллектор. В только что описанном режиме два эмиттера работают чуть выше отсечки и их эффективные величины α могут лежать в пределах от 0,01 до 0,1. Если увеличивать напряжение на зажимах, то когда

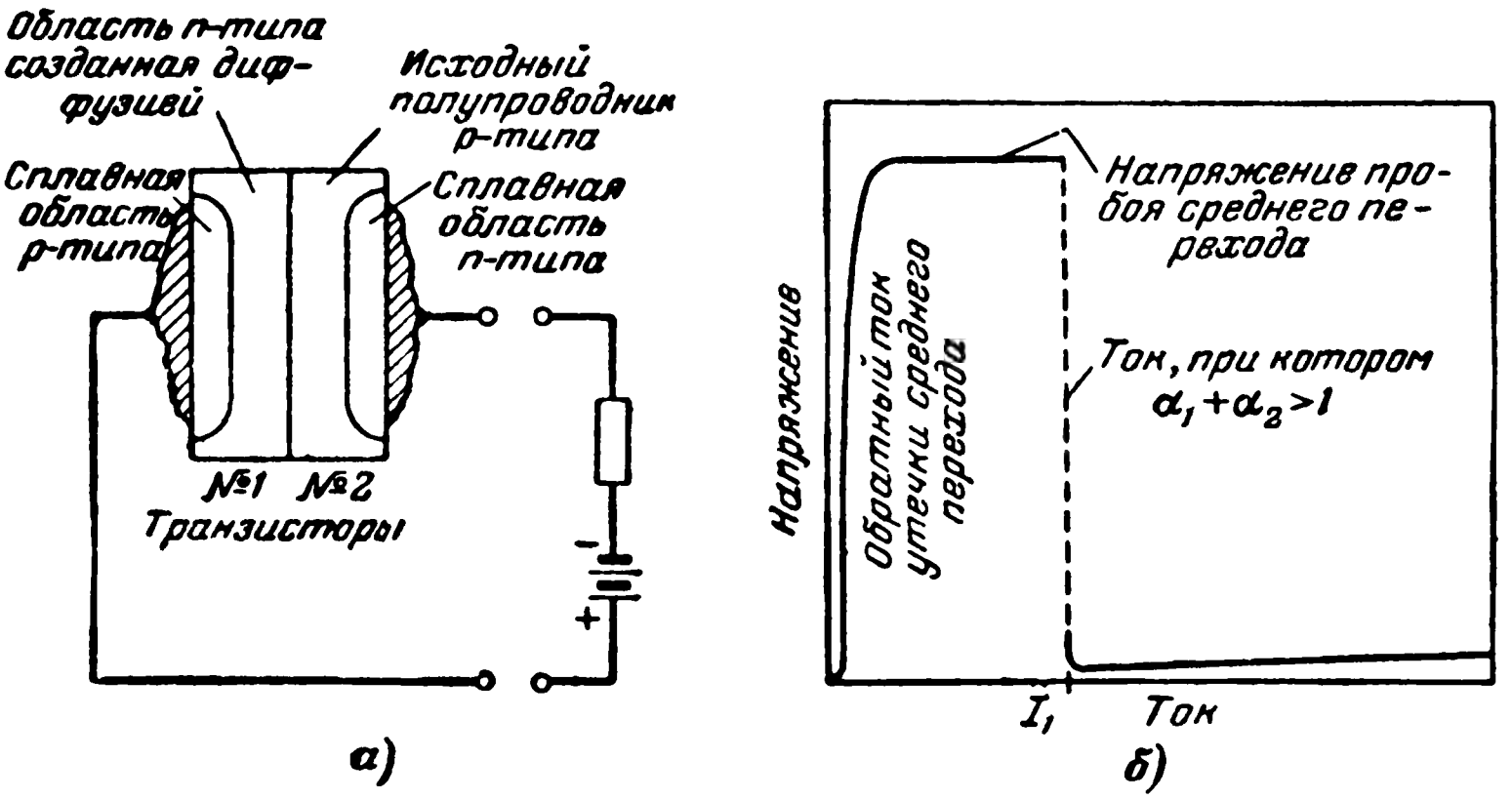


Рис 13-9 Структура *p-n-p-n* (*а*), служащая двухполюсником, обладает характеристикой с двумя устойчивыми положениями (*б*).

средний (коллекторный) переход достигнет напряжения пробоя $U_в$, ток начинает стремительно возрастать уже при небольшом повышении напряжения. При таком увеличении тока прямое напряжение обоих эмитерных переходов становится все бóльшим и бóльшим и α двух входящих в систему транзисторов приближается к единице. При некотором токе I_1, при котором сумма этих двух α становится больше единицы, возникает положительная обратная связь, и сопротивление между зажимами прибора внезапно падает до величины, равной сумме последовательно соединенных омических сопротивлений четырех слоев полупроводниковой структуры прибора. Для прибора, подобного показанному на рис. 13-9,*а*, имеющего две сплавные эмитерные области, сопротивление это может составлять всего 1—2 *ом*. Полная характеристика тогда будет такой же, как у прибора с двумя устойчивыми состояниями; на характеристи-

ке участки весьма высокого и весьма низкого сопротивления разделены областью отрицательного сопротивления. Такие приборы имеют очевидное применение в качестве ключей.

13-5. СИЛОВЫЕ ТРАНЗИСТОРЫ

Одной из трудных задач при разработке транзисторов является создание приборов, способных работать при мощностях в диапазоне от 1 до 10 *вт* и больше. В рабочем режиме обычные точечные транзисторы рассеивают мощность не более 200 *мвт*. У плоскостных

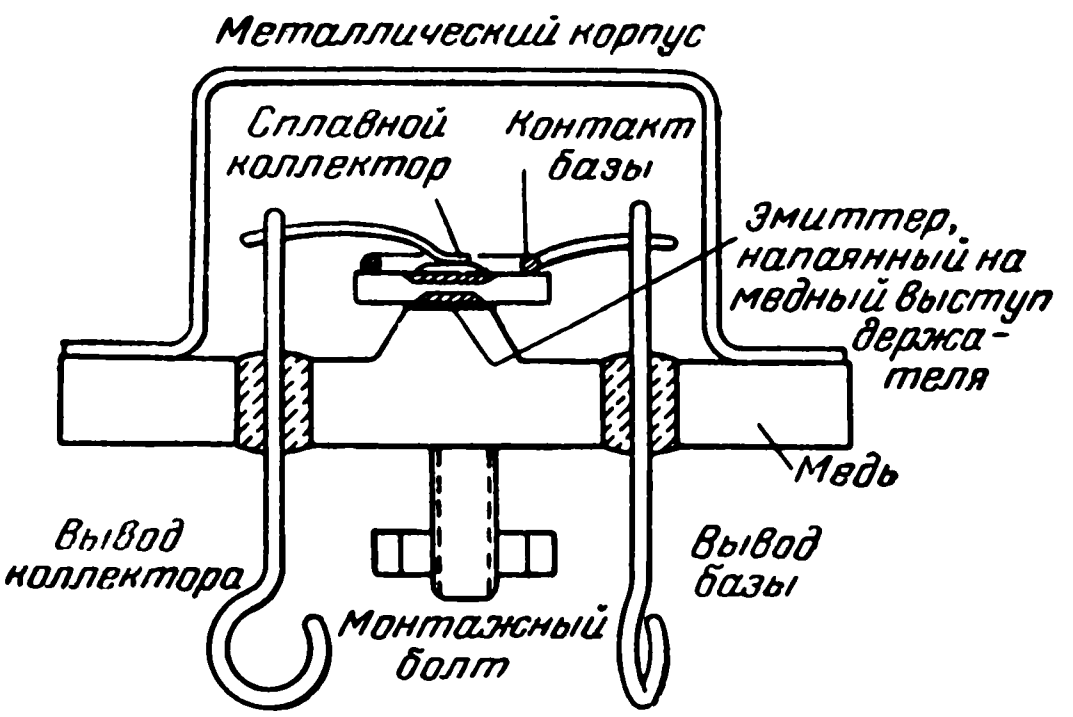


Рис. 13-10 Конструкция транзистора со сплавными переходами

транзисторов верхний предел допустимой рассеиваемой мощности ограничен (например, для таких специфических приборов, как тетроды) несколько меньшей величиной порядка 15—20 *мвт*. Во всех случаях это ограничение связано с трудностью отвода джоулева тепла. У германиевых приборов сопротивление *p-n* перехода при обратном напряжении уменьшается примерно на 8% на каждый градус прироста температуры. Таким образом, сравнительно небольшие изменения температуры могут вызвать значительные изменения входных и выходных полных сопротивлений транзистора в реальных цепях и соответственно токов, потребляемых транзистором при постоянных напряжениях источников смещения. Поэтому диапазон температур, в пределах которого можно ожидать неизменной работоспособности транзистора, ограничен температурой примерно на 50° С

выше комнатной (имеется в виду температура самого транзисторного элемента). Если тепло, выделяемое внутри транзисторного элемента на коллекторном переходе или на контакте, где в основном происходит тепловыделение, не может быть отведено, то приходится ограничивать допустимую рабочую мощность прибора.

Силовой транзистор — это прибор, физические размеры которого достаточно велики, чтобы пропустить токи требуемой величины [Л. 12], и в котором предусмотрены специальные меры для отвода тепла от транзисторного элемента [Л. 13]. На рис. 13-10 показан разрез силового плоскостного транзистора сплавного типа. Транзисторный элемент в виде пластинки германия имеет на обеих сторонах сплавные переходы и прикреплен к медному выступу, представляющему собой одно целое с теплоотводящей конструкцией корпуса. Чтобы пропускать через транзистор требуемый ток, размеры сплавных переходов в несколько раз больше, чем у обычных транзисторов. Так, у одного из разработанных силовых транзисторов типа *n-p-n* эмитерный переход имеет диаметр 3 *мм*, а коллекторный переход— 3,75 *мм*. Элемент может быть прикреплен к выступу основания эмитерной либо коллекторной стороной; обычно он укрепляется эмитерной стороной для удобства работы в схеме с общим эмиттером. Выводы коллектора и базы выполнены, как показано на рис. 13-10. Держатель снабжен болтом для обеспечения надежного теплового контакта с шасси или радиатором. Рабочие части прибора заключены в герметическую оболочку, заполненную сухим инертным газом.

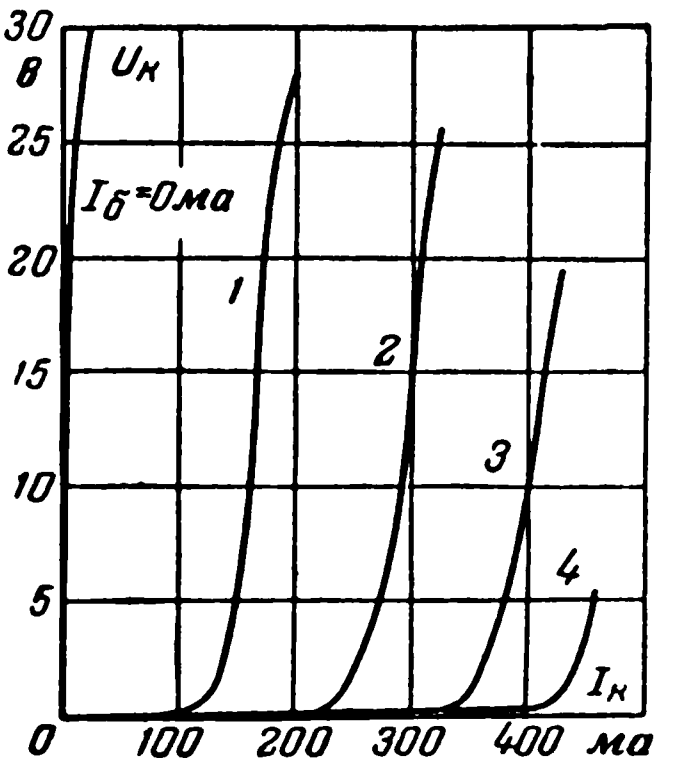


Рис. 13-11. Характеристики транзистора со сплавными переходами и структурой *n-p-n*, снятые при включении с общим эмиттером.

В серийных образцах сплавных транзисторов типа *n-p-n* перепад температур между корпусом и собственно транзисторным элементом может составлять при-

мерно 2° С на 1 *вт* рассеиваемой мощности. Если корпус транзистора привинчен к шасси обычных размеров и теплопроводности, то перепад температур между внешней средой и корпусом может составлять около 10° С на 1 *вт*. Описанный прибор рассчитан на мощность 2 *вт* для достаточной долговечности, хотя в некоторых случаях допустима его работа и при 4—5 *вт*. На рис. 13-11 приведено семейство коллекторных характеристик, снятое при включении такого транзистора по схеме с общим эмиттером. Изготовляются и другие типы сплавных транзисторов, рассчитанные на мощности до 25 *вт*, при условии обеспечения интенсивного теплоотвода с участков корпуса, предназначенных для этой цели.

ЗАДАЧИ

1. Сплавной силовой транзистор (рис 13-11) включен с общим эмиттером при напряжении источника питания коллектора 20 *в*, нагрузке в цепи коллектора 50 *ом* и входном токе смещения 1,5 *ма*. Определите графически:

а) коэффициент передачи тока между входом и выходом; б) мощность переменного тока, выделяемую в нагрузке при эффективном входном токе базы 0,5 *ма*; в) мощность, выделяемую током смещения в коллекторе; г) мощность, выделяемую постоянным током в нагрузке.

2. Постройте график концентрации электронов по длине стержневого транзистора *n*-типа, работающего с отсасывающим полем. Пусть теперь в стержень под инжекторный контакт введено облачко положительных дырок Постройте график концентрации дырок и электронов вдоль стержня для нескольких последовательных моментов времени. Постройте также график соответствующих распределений потенциалов. Как объяснить факт, что в то время как электроны дрейфуют в отсасывающем поле справа налево, облачко *избыточных* электронов, необходимых для нейтрализации пространственного заряда дырок, движется слева направо?

3. Какие характеристики получатся у канального транзистора при подаче на затвор различных *прямых* напряжений смещения? Начертите одну или две такие характеристики в качестве добавочных членов семейства на рис. 13-4.

4. Нарисуйте эквивалентную электрическую схему, изображающую различные термические сопротивления в тепловом потоке между коллекторным переходом силового транзистора и атмосферой, окружающей металлический корпус. Включите отдельно тепловые сопротивления германия, медного выступа корпуса, проволочного вывода коллектора и его штырька, стеклянных бусинковых впаев и переходное тепловое сопротивление между корпусом и окружающим воздухом Оцените приближенно относительные величины этих параметров.

ЛИТЕРАТУРА

1. Shockley W., Pearson G. L., Haynes J. R., Hole Injection in Germanium-Quantitative Studies and Filamentary Transistors, Bell System Technical Journal, 1949, vol. 28, p. 344.

2. Stuetzer O. M., A Crystal Amplifier with High Input Impedance, Proceedings of the Institute of Radio Engineers, 1950, vol. 38, p. 868.

3. Stuetzer O. M., Junction Fieldistors, Proceedings of the Institute of Radio Engineers, 1952, vol. 40, p. 1377.

4. Dacey G., Ross I. M., Unipolar Field-Effect Transistor, Proceedings of the Institute of Radio Engineers, 1953, vol. 41, p. 970.

5. Ross I. M., Field-Effect Transistors, Bell Laboratories Record, 1955, vol. 33, p. 167.

6. Miller S. L., Ebers J. J., Alloyed Junction Avalanche Transistors, Bell System Technical Journal, 1955, vol. 34, p. 883.

7. Miller S. L., Avalanche Breakdown in Germanium, Physical Review, 1955, vol. 99, p. 1234.

8. Shockley W., Sparks M., Teal G. K., p-n Junction Transistors, Physical Review, 1951, vol. 83, p. 151.

9. Ebers J. J., Four-Terminal p-n-p-n Transistors, Proceedings of the Institute of Radio Engineers, 1952, vol. 40, p. 1361.

10. Anderson A. E., Transistors in Switching Circuits Proceedings of the Institute of Radio Engineers, 1952, vol. 40; p. 1541.

11. Moll J. L., Tanenbaum M. Goldey J. M., Holonyak N., p-n-p-n Junction Transistors, Proceedings of the Institute of Radio Engineers, 1956, vol. 44, p. 1174.

12. Hall R. N., Power Rectifiers and Transistors, Proceedings of the Institute of Radio Engineers, 1952, vol. 40, p. 1512.

13. Roka E. G., Buck R. E., Reiland G. W., Development of Power Transistors, Proceedings of the Institute of Radio Engineers, 1954, vol. 42, p. 1247.

14. Полупроводниковые приборы и их применение, сборники статей под ред. Я. А. Федотова, вып. 1—7, Изд. «Советское радио», 1956.

ГЛАВА ЧЕТЫРНАДЦАТАЯ

ПРИМЕНЕНИЕ ТРАНЗИСТОРОВ

Несмотря на то, что транзисторы разработаны лишь около 10 лет назад, они уже широко используются в различной аппаратуре, заняли видное место в технике и промышленности и их роль продолжает расти. Транзисторы, например, уже прочно внедрены в приборах для тугоухих и карманных приемниках. По большей части они используются там, где существенна экономия в размерах аппаратуры и потребляемой мощности. Однако для переносного радиоприемника или телевизора,

которые с собой берут на пикник, желательны небольшие размеры, малый вес и небольшое потребление мощности. Точно так же и владельцы приборов для тугоухих заинтересованы в доведении этих качеств до возможно меньших пределов. В армейской аппаратуре и аппаратуре связи, где определенные габариты, вес и расход мощности часто являются обязательными требованиями, преимущества, обеспечиваемые транзисторами, могут позволить решить вопрос о том, продолжать ли дальнейшую разработку аппаратуры или отказаться от нее вовсе.

Литература по схемным вопросам применения транзисторов содержит сотни статей в технических журналах, не считая огромного количества всевозможного обзорного материала в популярных изданиях. Неудивительно, что дать исчерпывающий обзор всех случаев применения транзисторов не представляется возможным. Те частные случаи, которые описаны ниже, приведены лишь для иллюстрации наиболее типичных применений транзисторов [1].

14-1. УСИЛИТЕЛИ

Принципиальные схемы, приведенные на рис. 10-6, 10-8, 11-7 и 11-8, дают примеры включений транзисторов в усилительные схемы с общей базой, общим эмиттером и общим коллектором. Эти схемы являются основными и показаны в идеализированном виде, т. е. в них отсутствуют добавочные элементы, которые включаются для достижения желаемых результатов, например автоматического смещения, ограничения полосы пропускания, стабилизации за счет применения обратной связи и т. д. Однако эти схемы можно считать типовыми для применения транзисторов в усилителях приборов для тугоухих, в каскадах усиления высокой промежуточной и низкой частот радиоприемников, в телефонных промежуточных усилителях и многих других специальных усилителях.

На рис. 14-1 показана схема низкочастотного усилителя класса А для выходного каскада радиоприемника или граммофонного проигрывателя Здесь использованы два транзистора в двухтактной схеме с общей базой,

---

[1] Более подробный перечень применения транзисторов см. в литературе к гл. 11.

чтобы получить неискаженную мощность, бóльшую той, которую может дать схема на одном транзисторе. Сопротивления $R_в$ (которые по отношению к транзисторам являются внешними и которые не следует смешивать с собственным параметром $r_б$ эквивалентной схемы транзисторов) используются для автоматического смещения баз. Через сопротивления $r_а$, включенные между базами и источником питания коллектора, протекает до-

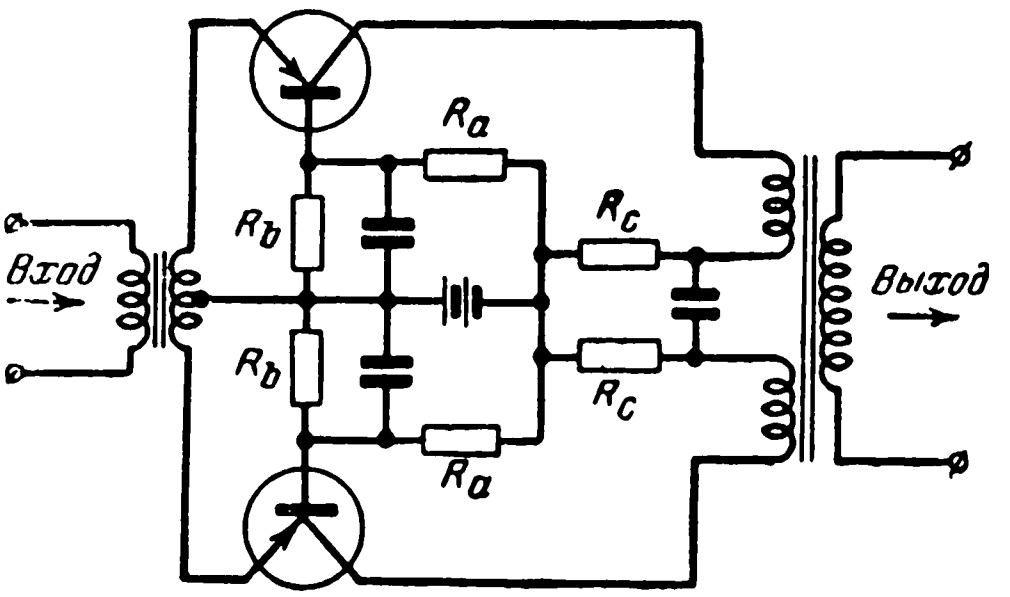


Рис 14-1 Два транзистора в двухтактном усилителе низкой частоты по схеме с заземленной базой, работающем в режиме класса А

бавочный ток (он течет также через сопротивления $R_в$), увеличивающий смещения баз. Сопротивления $R_в$ не нужно брать слишком большими, так как это приведет к нестабильности схемы, потому что $R_в$ является элементом положительной обратной связи. Сопротивления $R_в$ служат для балансировки выходов двух транзисторов, которые в общем случае не являются абсолютно идентичными. Эти сопротивления, включенные, как показано на схеме, не расходуют мощности переменного тока, так как напряжение переменного тока на первичной обмотке выходного трансформатора равно нулю в средней точке обмотки.

Еще бóльшую неискажаемую выходную мощность можно получить, перейдя на работу в классе АВ или В. В этих случаях сопротивления автоматического смещения уменьшаются или полностью исключаются. В последнем крайнем случае рабочая точка покоя эмиттера по постоянному току соответствует примерно отсечке тока эмиттера и работе в классе В.

Практическая схема включения транзистора с общим эмиттером показана на рис. 14-2. Этот усилитель рассчитан на использование тетрода в качестве миниатюрного промежуточного усилителя для коаксиальной линии связи [Л. 1]. Транзистор имеет трансформаторную связь как с входом, так и с выходом. Схема содержит несколько интересных решений для получения смещения и стабилизации режима. Постоянное рабочее смещение между

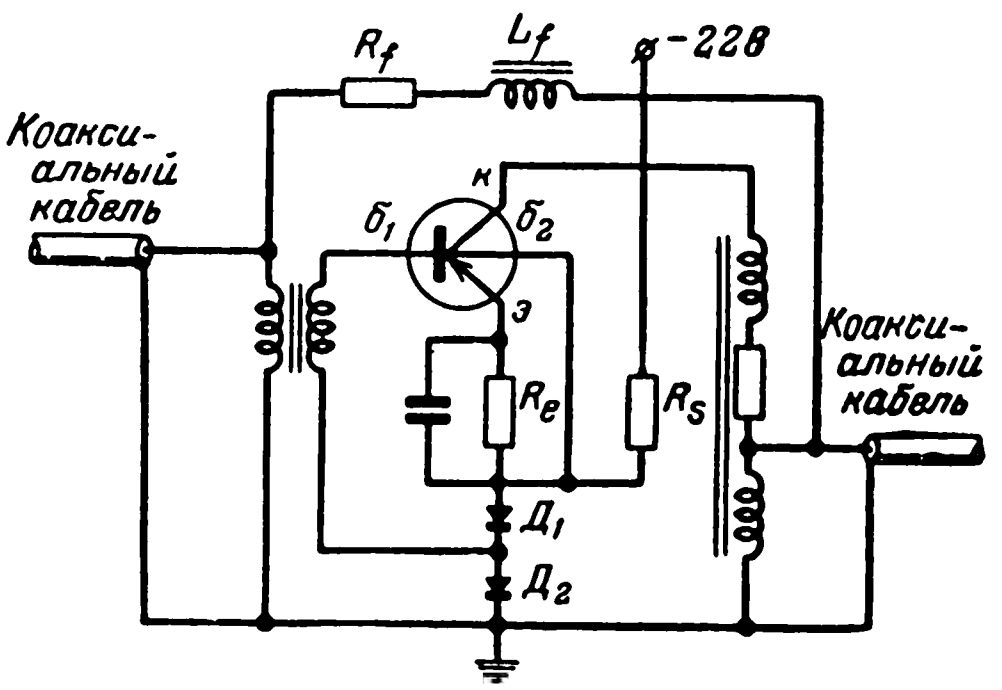


Рис 14-2 Усилитель на тетроде для коаксиальной линии связи, обеспечивающий усиление 22 *дб* в пределах полосы частот шириной 12 *Мгц*.

эмиттером и двумя выводами базы снимается с делителя напряжения, образованного последовательно включенными сопротивлением R_e и двумя диодами: D_1 и D_2. В качестве последних применены кремниевые сплавные диоды, каждый из которых работает при напряжении выше напряжения обратного пробоя в диапазоне токов, в котором изменение тока практически не вызывает изменения напряжения на выводах диодов. Сопротивление R_c используется в качестве элемента, стабилизирующего положение рабочей точки эмиттера по постоянному току при температурных изменениях сопротивления эмиттера и, следовательно, тока эмиттера. Стабилизация полного коэффициента усиления достигается с помощью отрицательной обратной связи с выхода на вход через R_f и L_f. Усилитель потребляет 5 *ма* и питается от одного источника напряжением 22 *в*. Схема может быть легко приспособлена для питания постоянным током, подаваемым по

центральной жиле коаксиального кабеля. Усилитель обеспечивает усиление 22±0,5 *дб* в полосе частот от 350 *кгц* до 12 *Мгц*. Весь усилитель имеет длину примерно 38 *мм* и диаметр около 3 *мм*.

14-2. ГЕНЕРАТОРЫ И МУЛЬТИВИБРАТОРЫ

Генераторы, разработанные с использованием транзисторов в качестве активных элементов схем, широко распространены и приспособлены к транзисторам определенных параметров. На рис. 14-3 показаны две схемы синусоидальных генераторов. В первом из них [Л. 2]

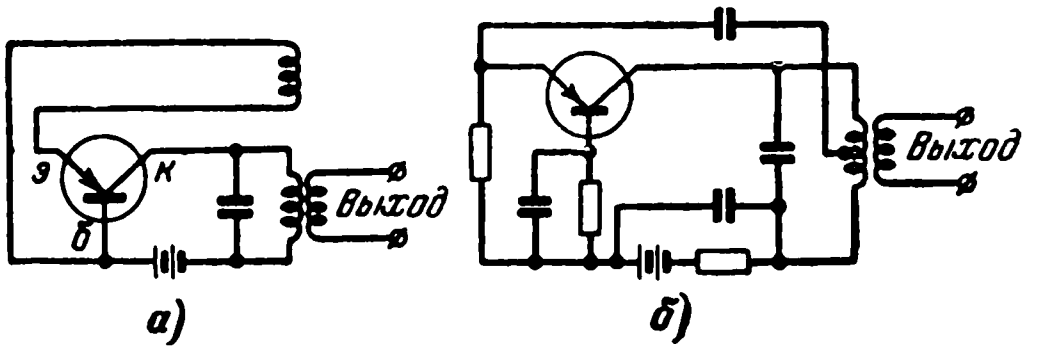


Рис. 14-3 Схемы генераторов синусоидальных колебаний на транзисторах

применена индуктивная обратная связь между коллектором и эмиттером, во втором — емкостная связь с отвода катушки контура.

Хорошей демонстрацией возможностей транзисторов для уменьшения размеров переносной аппаратуры является передатчик с частотной модуляцией [Л. 3], принципиальная схема которого приведена на рис. 14-4. Единственный транзистор служит как для генерирования высокой частоты, так и для частотной модуляции. В этой схеме применен специально изготовленный точечный транзистор, имеющий высокую частоту отсечки α. Даже в том случае, когда предельная частота f_α для выбранного транзистора равна примерно 20 *Мгц*, цепь имеет достаточное усиление, чтобы обеспечить генерирование на частоте примерно 104 *Мгц*, позволяющей еще излучать полезный сигнал.

Высокочастотная часть схемы представляет собой обычный генератор с обратной связью через межэлектродные емкости. Частотная модуляция осуществляется перемещением мгновенной рабочей точки эмиттера низкочастотным сигналом, поступающим от микрофона. Это перемещение рабочей точки эмиттера приводит к соот-

ветствующему изменению усиления высокой частоты. Поскольку при работе на высокой частоте транзистор находится в режиме, соответствующем рабочей точке, находящейся далеко внизу на склоне частотной характеристики (рис. 12-4), указанные колебания вызывают сдвиг фазы генерируемого сигнала. В такой схеме модуляции получается фазово-модулированный сигнал, который с точки зрения его характера и способа приема не отличается от частотно-модулированного. В этом передатчике катушка контура является также и передающей антенной. Схема питается батарейкой, предназначенной для миниатюрного фонарика в карандаше, и вместе с микрофоном и питанием занимает объем около 55 *см*^{3}.

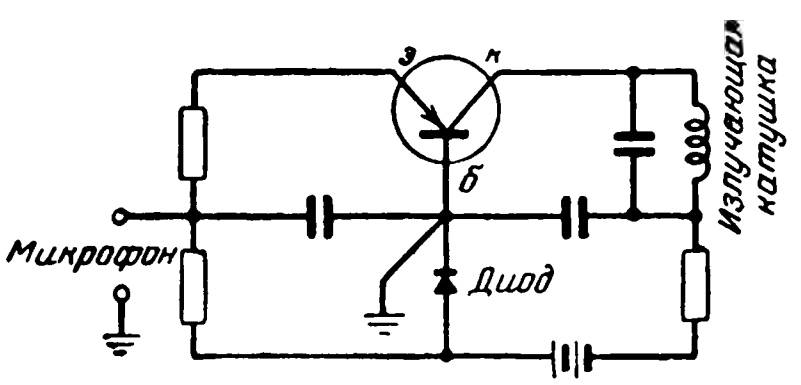


Рис 14-4. Частотно-модулируемый передатчик на частоте 104 *Мгц*.

При дальнейшей разработке возможно еще большее уменьшение объема. Катушка контура (шесть витков диаметром 12 *мм* и такой же длины) является малоэффективным излучателем на частоте 104 *Мгц*, но даже в этом случае сигналы передатчика могут быть приняты стандартным ЧМ-приемником на расстоянии в несколько сотен метров.

На рис. 14-5 показана схема самовозбуждающегося генератора напряжения прямоугольной и пилообразной форм [Л. 4] на точечном транзисторе, работающего следующим образом. Когда ключ в цепи эмиттера разомкнут, база транзистора отрицательна за счет падения напряжения IR обратного тока утечки коллектора на сопротивлении базы R_b. При замыкании ключа напряжение эмиттера вначале равно нулю, т. е. эмиттер положителен относительно базы. Поэтому начинает идти ток эмиттера; ток коллектора возрастает, база становится еще более отрицательной, и полное сопротивление между эмиттером и коллектором транзистора падает до очень малой величины. При этом конденсатор заряжается через R_d. Этот (самоподдерживающийся) лавинный процесс продолжается до тех пор, пока разность напряжений между базой и коллектором не станет слишком малой для поддержания падения напряжения

IR_b бóльшим, чем отрицательное напряжение, которое уже имелось на конденсаторе со стороны эмиттера. Вследствие этого схема отпирается, напряжение на базе принимает свою первоначальную величину, а напряжение эмиттера медленно поднимается по мере стекания заряда через R_e. Когда эмиттер снова становится менее отрицательным, чем база, весь цикл работы повторяется. Напряжение на емкости имеет приблизительно пилообразную форму, тогда как форма импульсов тока через R_d приблизительно прямоугольна. Длительность генерируемого импульса определяется величиной R_dC_e, а частота повторения — величиной R_eC_e.

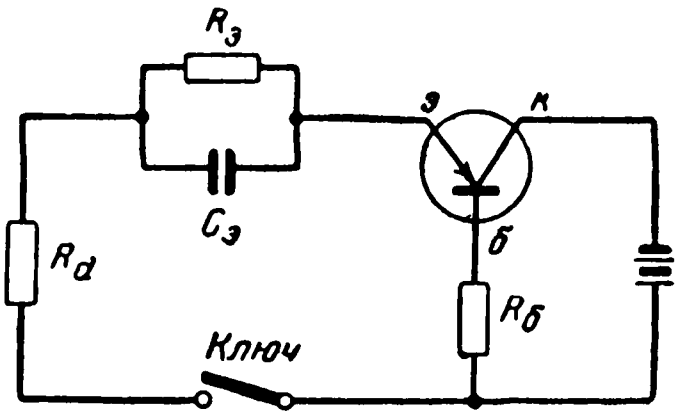


Рис. 14-5. Самовозбуждающийся генератор пилообразных и прямоугольных импульсов на транзисторе.

14-3. ИМПУЛЬСНЫЕ СХЕМЫ НА ТРАНЗИСТОРАХ

Можно легко видоизменить схему на рис. 14-5 так, что она превратится в полупериодный мультивибратор (рис. 14-6). Батарея смещения включена в цепь эмит-

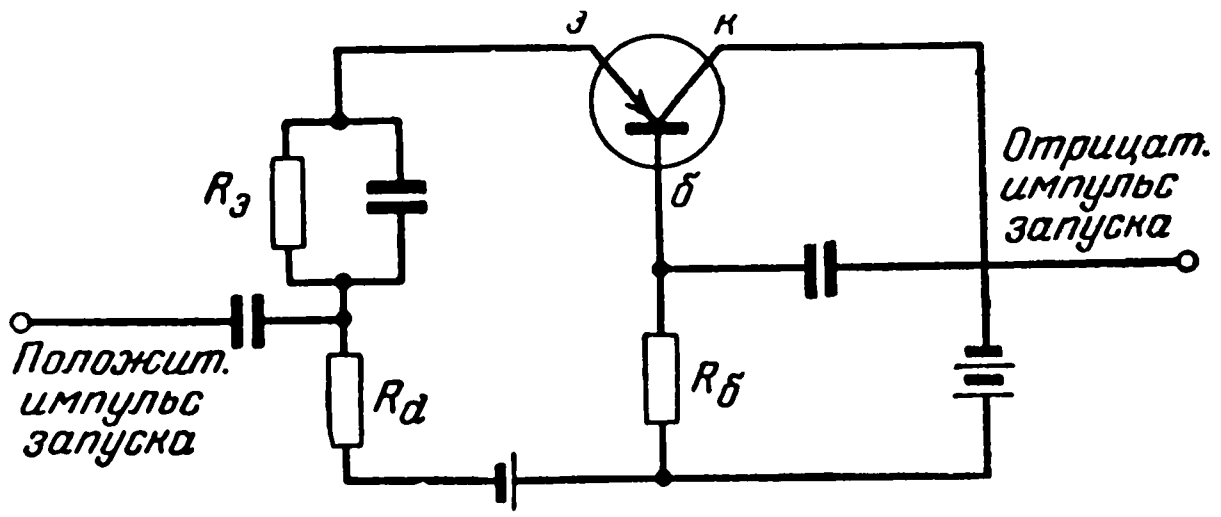


Рис 14-6 Введение батареи смещения в схему, показанную на рис 14-5, превращает ее в ждущий одностабильный мультивибратор.

тера для того, чтобы в установившемся режиме эмиттер оставался отрицательным относительно базы, обеспечивая, таким образом, устойчивое состояние запертого транзистора. Рабочий ход может быть начат подачей отрицательного импульса на базу или положительного импульса на эмиттер, причем их амплитуда должна быть такова, чтобы на один момент база оказалась от-

рицательнее эмиттера. В результате схема совершает рабочий ход, восстанавливается и снова принимает устойчивое положение при высоком сопротивлении и малом токе транзистора, сохраняемое до поступления следующего запускающего импульса. Такая схема может служить генератором импульсов. Запуск может быть произведен импульсами малой амплитуды от источника с вы-

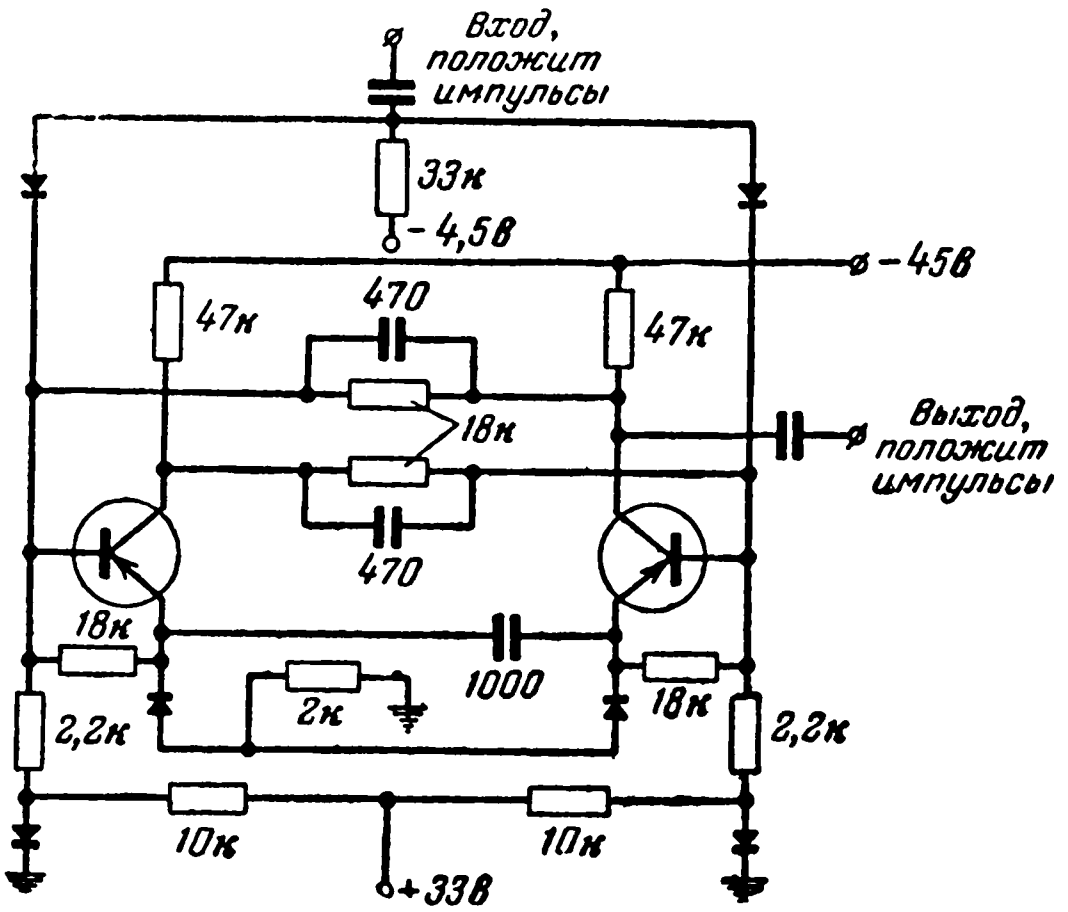


Рис 14-7 Триггерная цепь с двумя устойчивыми состояниями, выдающая выходной импульс при каждом втором входном импульсе.

соким внутренним сопротивлением; генерируются импульсы значительно большей амплитуды при сравнительно низком выходном сопротивлении. Кроме того, генерируемые импульсы имеют всегда одну и ту же форму, амплитуду и длительность независимо от характера запускающих импульсов, лишь бы они были в состоянии запустить схему. Такая особенность бывает полезна при различных видах селекции импульсов; уровень селекции может быть установлен регулировкой чувствительности схемы к запускающему напряжению.

Описанная схема является основой триггера на транзисторах (рис. 14-7) [Л. 5]. Она является примером бистабильной схемы, способной сохранять одно из двух возможных устойчивых состояний, в любое из которых она может быть переключена по желанию. Эти два состояния являются взаимно исключающими в том смыс-

ле, что установление одного состояния автоматически выводит схему из другого. Триггерные схемы служат базовыми элементами электронных счетных машин, схем накопления информации и других логических схем для распределения и обработки данных.

Подробный разбор работы всей схемы на рис. 14-7 затруднен, но можно отметить одну особенность, а именно связь коллектора каждого транзистора с базой другого. Таким образом, если вначале один транзистор находится в отпертом состоянии и другой транзистор внезапно приводится внешним импульсом в такое же состояние, то на базу первого транзистора будет подано менее отрицательное напряжение, вследствие чего он будет переведен в запертое состояние. Повторение запускающего импульса заставит транзисторы снова поменяться ролями. Входная схема на диодах служит для подачи запускающего импульса на тот транзистор, который в настоящий момент заперт, и для запирания другого транзистора.

Существенная особенность работы триггера заключается в том, что для восстановления первоначального состояния схемы требуются два последовательных импульса. Схема, показанная на рис. 14-7, выдает один положительный выходной импульс на каждый второй положительный входной импульс. Схема, таким образом, может быть использована как числовой регистр в двоичном счетчике. Если несколько таких триггеров соединить последовательно, подав выход одного каскада на вход другого, то это устройство может служить для счета импульсов, причем состояние нескольких каскадов будет зависеть от числа импульсов, поступивших на первый каскад. Двоичные логические операции, осуществляемые с помощью моностабильных и бистабильных схем охватывают огромную область применений и многократно описаны в литературе [Л. 6].

14-4. СЕЛЕКТОРНЫЕ И КЛЮЧЕВЫЕ СХЕМЫ С ТРАНЗИСТОРАМИ

На рис. 14-8 показана селекторная схема с транзистором. Она допускает прохождение сигнального импульса только при подаче отпирающего напряжения на вход управления. Напряжение эмиттера обычно поддерживается ниже отсечки, так что сигнал не пропускается на

выход из-за обратных сопротивлений эмиттера и коллектора. Если на вход управления приходит напряжение селекторного импульса и если его величина достаточна, чтобы получить на эмиттере прямое смещение, при котором транзистор начинает работу, то импульс сигнала не только проходит на выход, но и усиливается. Эта схема в варианте, показанном на рис. 14-8, пригодна для отбора положительных импульсных сигналов с ам-

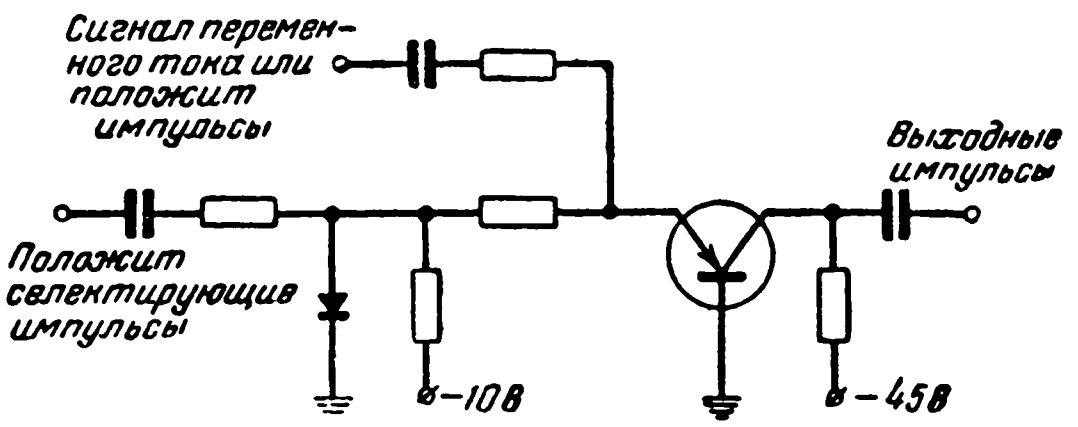


Рис. 14-8 Схема совпадения (импульсный селектор), не только выполняющая селекцию, но и усиливающая выделенный сигнал

плитудой до нескольких вольт или для отбора сигналов переменного тока малой амплитуды.

Интересно сравнить эту схему с диодной селекторной схемой типа *И* (рис. 7-15). С функциональной точки зрения обе схемы подобны, но схема на транзисторах имеет то преимущество, что она усиливает сигнал, компенсируя потери, имеющие место в схеме.

На рис. 14-9 показана типовая схема включения плоскостного транзистора в качестве ключа с общим эмиттером. В зависимости от того, подано ли на переход между эмиттером и базой обратное или прямое смещение, ключ будет разомкнут или замкнут. В некоторых случаях бывает желательно подавать постоянное напряжение смещения надлежащего знака, чтобы держать ключ в замкнутом или разомкнутом состоянии. В других случаях возможность самоблокирования ключа в замкнутом или разомкнутом состоянии, в которое он был приведен с помощью соответствующих управляющих импульсов, обеспечивается прибавлением ряда схемных элементов. Такое свойство самоблокирования имеет место у ключа на лавинном транзисторе (рис. 13-6,*а*) или у ключа на транзисторе типа *p-n-p-n* (рис. 13-9,*а*).

Часто бывает важно знать быстрогу срабатывания ключа. В большинстве ключей на транзисторах длительности процессов замыкания и размыкания мало зависят от емкости перехода и в большей степени определяются временем накопления неосновных носителей в активной области базы. Когда такой ключ замкнут, через отпертый транзистор между эмиттером и коллектором проходит сравнительно большой ток. При этом область базы, через которую этот ток проходит в виде потока неосновных носителей, наполняется этими частицами с концентрацией, намного превышающей ту, которая имеет место при запертом транзисторе. Как показано на рис. 14-10, для установления высокой проводимости транзистора после подачи отпирающего смещения требуется время.

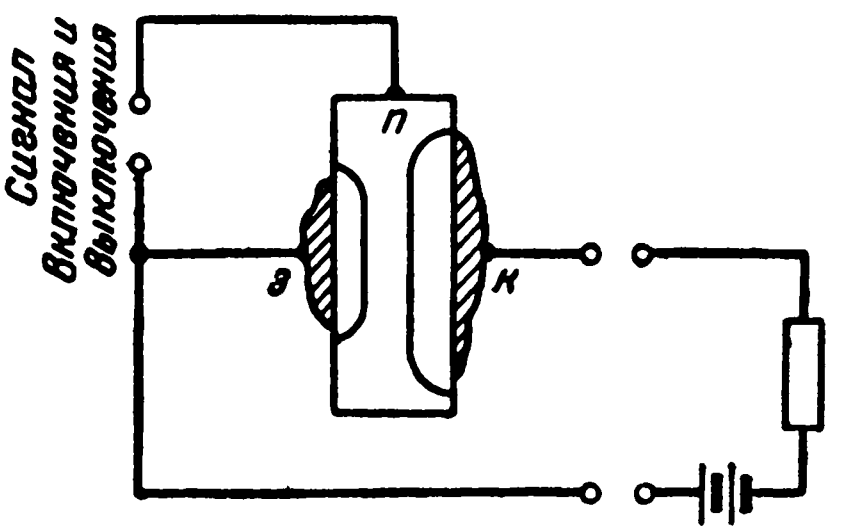


Рис 14-9. Типовая схема наиболее часто применяемого транзисторного ключа с общим эмиттером.

Если затем знак этого смещения мгновенно меняется так, чтобы снова запереть гранзистор, то восстановление запертого состояния снова происходит с задержкой. Она соответствует времени, в течение которого заряд избыточных неосновных носителей, имевшийся в активной области базы в момент перемены знака смещения на эмиттере, исчезает из этой области. После того как это произойдет, частично путем диффузии к коллекторному переходу слева направо, частично путем диффузии справа налево к эмитерному переходу и частично путем рекомбинации в самом слое базы, через оба перехода продолжают идти токи больше тех, которые соответствуют участкам насыщения на характеристиках при обратном напряжении. Пока это исчезновение заряда не закончится, ключ продолжает проводить. Как указано на рис. 14-10, в первый момент после запирания диффузия происходит столь стремительно, что ток остается в основном неизменным и определяется напряжением питания и общим сопротивлением цепи. В течение последующей части процесса запирания ток спа-

дает и постепенно устанавливается на уровне тока насыщения коллекторного перехода при обратном напряжении.

Физические процессы нарастания тока при отпирании и спаде тока при запирании транзистора можно представить аналитически [Л. 10] в виде выражений, связывающих время установления, с другими параметрами транзистора и его внешней цепи. Как можно было ожидать, длительности нарастания, сохранения и спада

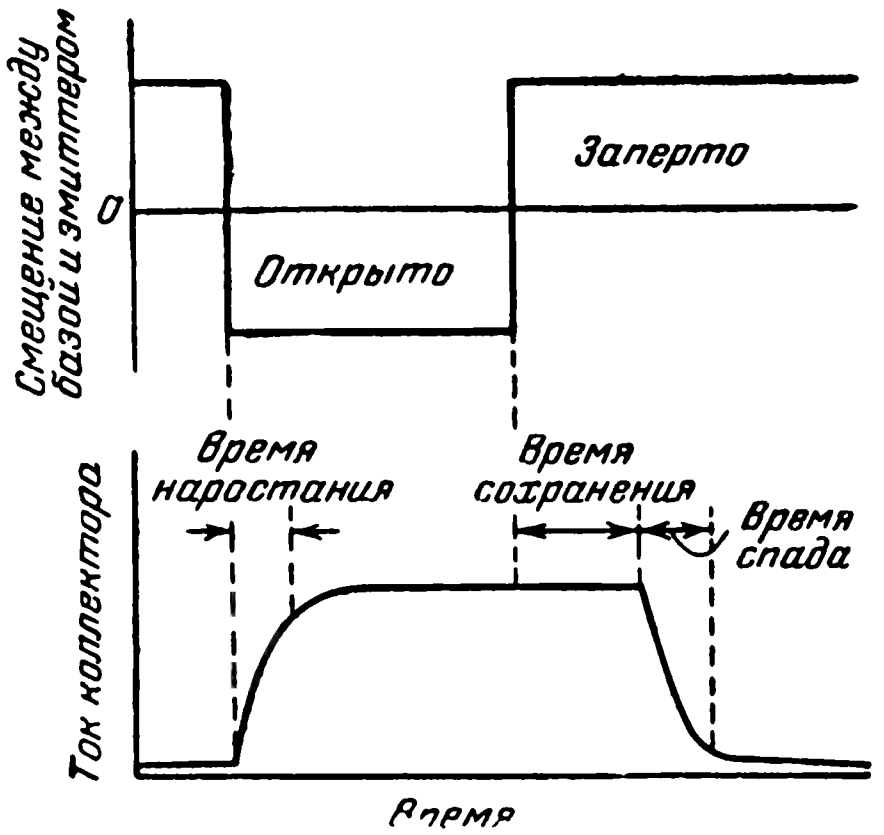


Рис 14-10 При запирании или отпирании ключевого транзистора новый режим устанавливается лишь после короткой задержки

связаны с верхней граничной частотой транзистора, поскольку эта частота подобным же образом определяется процессами диффузии неосновных носителей в слое базы. Эти времена могут быть уменьшены, если сделать слой базы более тонким. Длительность сохранения также зависит от тока, протекавшего непосредственно перед моментом запирания транзистора, и от величины запирающего напряжения, поданного на базу. Типичная величина времени нарастания тока для обычных сплавных транзисторов при их работе в ключевых схемах может достигать 0,1—0,2 *мксек*, время сохранения—около 0,5 *мксек* и время спада—около 0,1—0,2 *мксек*. При расчете быстродействующих схем разработчик должен знать об этих ограничениях и вводить соответствующие поправки при определении момента коммутации.

ЗАДАЧА

1. Импульсная схема совпадения типа НЕТ, И представляет собой схему, которая выдает выходной импульс, если на один из двух входов поступает импульс; но если на оба входа поступают одновременно два импульса, на выходе схемы импульса не возникает. Операция НЕТ, И играет роль при выполнении двоичного сложения, при котором 0+0=0; 1+0=1; 0+1=1; 1+1=0+перенос разряда. Составьте простую схему типа НЕТ, И на транзисторах.

ЛИТЕРАТУРА

1. Straube H. M., A Miniature Transistor Amplifier, Bell Laboratories Record, 1955, vol. 33, p. 237.

2. Ryder R. M., Kircher R. J., Some Circuit Aspects of the Transistor, Bell System Technical Journal, 1949, vol. 28, p. 367.

3. Thomas D. V., A Point-Contact Transistor FM VHF Transmitter, Transactions of the Institute of Radio Engineers, Professional Group on Electron Devices, 1954, vol. ED-1, p. 43.

4. Anderson A. E., Transistors in Switching Circuits, Proceedings of the Institute of Radio Engineers, 1952, vol. 40, p. 1541.

5. Trent R. L., Binary Counter Using Two Transistors, Electronics July 1952, vol. 25, p. 100.

6. Proceedings of the Institute of Radio Engineers, 1952, vol. 40.

7. Moll J. L., Large-Signal Transient Response of Junction Transistors, Proceerings of the Institute of Radio Engineers, 1954, vol. 42, p. 1773.

8. Полупроводниковые приборы и их применение, сборники статей под редакцией Я. А. Федотова, вып. 1—7, Изд. «Советское радио», с 1956 г.

9. Полупроводниковые приборы с отрицательным сопротивлением, сборник статей под редакцией С. А. Гаряинова, Госэнергоиздат, 1962.

10. Зборовский М. И., Научно-техническая конференция по применению полупроводников в приборостроении, «Вестник машиностроения», 1957, № 9, стр. 72—74.

11. Полупроводниковые триоды и их применение, под ред. А. В. Красилова, Госэнергоиздат, 1957.

ЧАСТЬ ВТОРАЯ

ФИЗИКА ТВЕРДЫХ ТЕЛ И ЕЕ ПРИМЕНЕНИЕ К КОНСТРУИРОВАНИЮ ПОЛУПРОВОДНИКОВЫХ ПРИБОРОВ

ГЛАВА ПЯТНАДЦАТАЯ

ЗОННАЯ ТЕОРИЯ ТВЕРДЫХ ТЕЛ

В предшествующих главах мы рассмотрели полупроводники и полупроводниковые приборы с феноменологической точки зрения и познакомились с теми свойствами полупроводников, которые можно использовать в полупроводниковых приборах.

Теперь мы изучим сами полупроводники. Мы рассмотрим детально поведение атомов, молекул, электронов и положительно заряженных дырок. Мы воспользуемся при этом уже известными нам сведениями и столкнемся с новыми для нас идеями в физике, которые придется пока принять без доказательств. Так, мы создадим модель, которая объяснит наблюдавшиеся нами особенности работы полупроводников во всех случаях их температурной зависимости, выпрямления, действия как транзисторов, фотоэлектрических свойств, эффекта Холла и термоэлектрического эффекта. Основой для понимания электронных свойств твердых веществ является знание зонной теории электронов. Она может быть создана на основе какой-либо из двух возможных точек зрения — атомистической или исходящей из коллективного поведения электронов. В атомистическом приближении следует сначала изучить поведение электронов в изолированных атомах и молекулах. Затем следует сгруппировать большие количества отдельных атомов или молекул для образования твердого кристалла и посмотреть, какие изменения в поведении электро-

нов произойдут в результате этого группирования При рассмотрении этого вопроса, исходя из приближения сгруппированных электронов, следует рассмотреть движение электронов в электронном газе, заключенном внутри твердого кристалла. Затем накладываются квантовые ограничения, вводятся искажения периодичности кристаллической решетки и исследуется влияние этих уточнений на поведение электронного газа. В двух следующих ниже параграфах приведено элементарное изложение способов решения задачи на основе обеих точек зрения.

15-1. ЗОННАЯ ТЕОРИЯ С ТОЧКИ ЗРЕНИЯ АТОМИСТИЧЕСКОГО ПРИБЛИЖЕНИЯ

При построении этой теории мы начнем с модели водородного атома Бора, которая представляет собой ядро атома в виде бесконечно малой частицы, несущей положительный заряд e, равный по величине, но противо-

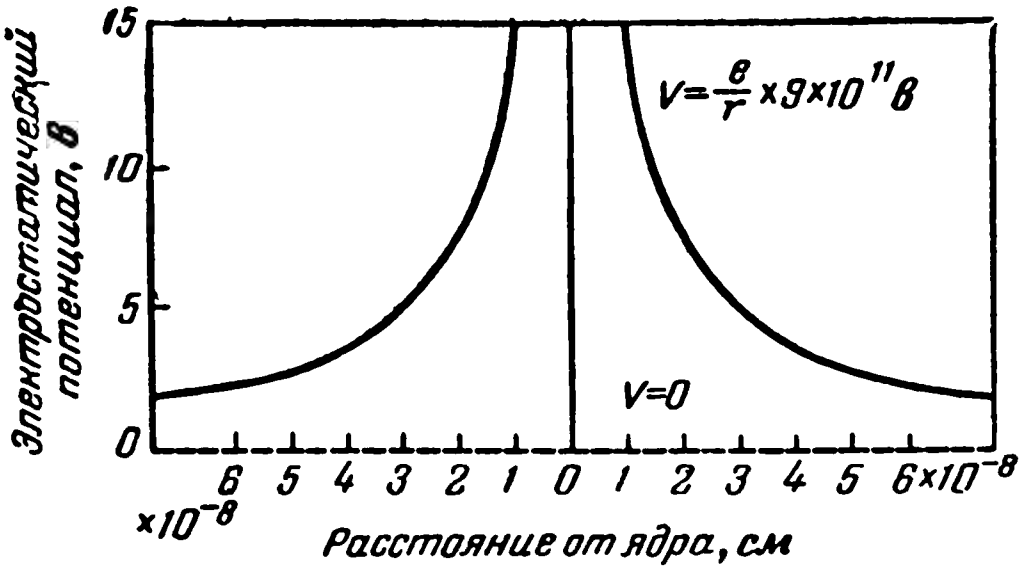


Рис 15-1 Электростатический потенциал вблизи водородного ядра изображается гиперболой.

положный по знаку заряду электрона — e. Электрон, принадлежащий данному атому, вращается вокруг ядра по орбите, как планета в миниатюрной солнечной системе. Если мы уберем этот электрон и построим график электростатического потенциала вблизи изолированного ядра, мы получим зависимость, показанную на рис. 15-1. Нулевой потенциал изображен горизонтальной пунктирной линией, соответствующей потенциалу на бесконечном расстоянии от ядра. Кривая на рис. 15-1

есть просто график потенциала по уравнению $V=(e/r)\ 9\cdot 10^{11}$ *в*, где e — заряд ядра ($1{,}6\cdot 10^{-19}$ *к*) и r — расстояние от ядра в сантиметрах.

Если теперь внести электрон куда-нибудь в близлежащую к ядру область, то эта система произведет работу, так как ядро и электрон противоположно заряжены и энергия получившейся в результате системы будет отрицательной. Иначе говоря, электрон обладает отрицательной энергией, если он находится в области с по-положительным потенциалом При обсуждении атомных процессов удобно использовать в качестве единицы энер-

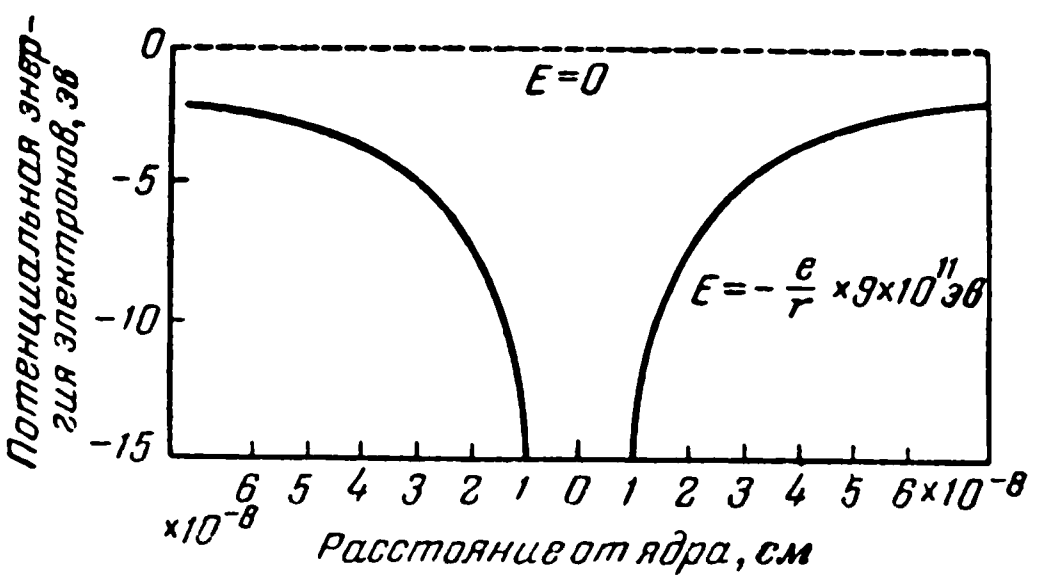


Рис 15-2 Потенциальная энергия электрона вблизи ядра водородного атома изображена на графике гиперболой Сравните его с графиком на рис 15-1.

гии электронвольт. 1 *эв* является кинетической энергией одного электрона, который, начав движение из состояния покоя, проходит разность потенциалов 1 *в*. Если потенциал в определенном месте пространства равен V *в*, то электрон в этом месте обладает потенциальной энергией — V электронвольт или — eV джоулей и чтобы изобразить график потенциальной энергии электрона в функции расстояния от ядра водорода, надо просто повернуть кривую на рис. 15-1 вокруг оси расстояний и дать название ординатам «Энергия электронов, *эв*», как это сделано на рис. 15-2.

Если мы теперь спросим, какие различные положения может занимать орбита электрона в пределах потенциальной ямы, классическая механика дает нам ответ, что электрон может занимать любую из бесконечного множества орбит, круговых или эллиптических, и

может обладать любой из бесконечного множества возможных энергий. При таком представлении движение электрона было бы очень похоже на движение шарика, катящегося без трения по внутренним стенкам воронки. На рис. 15-3 показана одна из таких орбит, слегка наклоненная для создания перспективы. Потенциальная энергия (в электронвольтах) и электрона на этой орбите определяется графически расстоянием по вертикали от горизонтальной пунктирной линии вниз до уровня орбиты. Положительная кинетическая энергия электро-

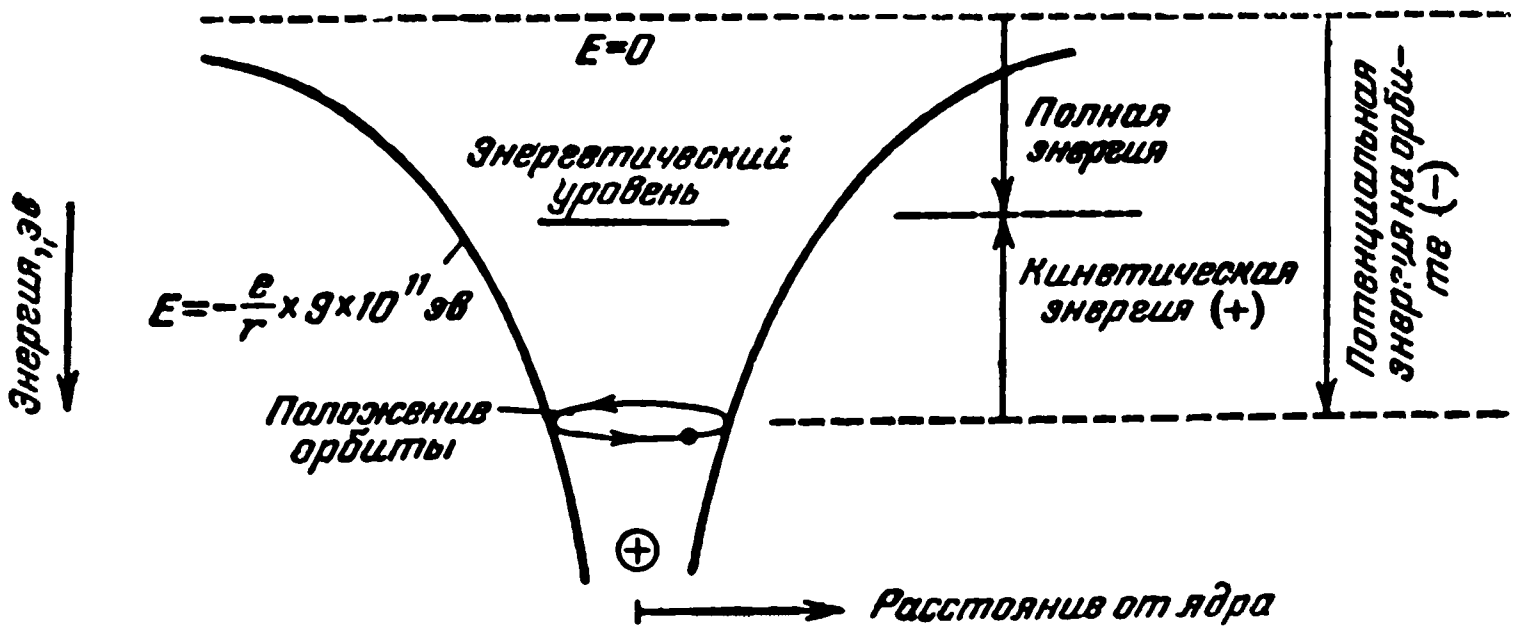


Рис 15-3 На графике энергии электрона вблизи ядра водородного атома показаны орбита электрона и соответствующий ей уровень полной энергии системы

на изображается более короткой стрелкой, направленной вверх от этого уровня. Для всех систем, силы в которых подчиняются закону обратного квадрата расстояния от центра (солнечная система и атомы), величина положительной кинетической энергии частицы на стабильной круговой орбите всегда равна половине отрицательной потенциальной энергии. Поэтому длина стрелки, обозначающей кинетическую энергию, такая, какая показана на рисунке, а полная энергия системы дается направленной вниз стрелкой, обозначенной «Полная энергия». Условно полагают, что электрон занимает «энергетический уровень», изображенный горизонтальной линией с надписью «энергетический уровень». Для того чтобы ионизовать атом, сорвав электрон с орбиты и удалив его на бесконечное расстояние от ядра, необходимо затратить энергию, изображаемую длиной стрелки, представляющей полную энергию.

Это рассуждение может быть в равной мере применено и к случаю эллиптической орбиты, для которой средняя кинетическая энергия равна половине средней потенциальной энергии (при усреднении за один полный обход электрона по орбите). Полная энергия электрона на такой орбите остается постоянной и может быть изображена, как и раньше, энергетическим уровнем — горизонтальной линией, проведенной на соответствующем расстоянии ниже пунктирной линии $E=0$.

Из этого бесконечного множества возможных классических орбит и соответствующих им энергий согласно законам квантовой механики электрон может занимать какой-либо из ограниченного числа «разрешенных» энергетических уровней в пределах потенциальной воронки. Эти ограничения, налагаемые квантовой механикой, вступают в действие каждый раз, когда перемещения электрона ограничены воздействием внешних сил, например при захвате внутри потенциальной ямы ядра атома или же в области, ограниченной поверхностями граней кристалла Очень упрощенную картину того, что происходит в результате этого ограничения, можно представить следующим образом.

Поведение электрона описывается с помощью его «волновой функции». Если Вы спросите, что это за волна, Вы не получите ответа. Это не водяная волна, не звуковая волна, не электромагнитная или же какая-либо другая из известных Вам волн. Это всего лишь колебательное свойство пространства, в котором находится электрон. Одни предпочитают считать, что электронная волновая функция является чисто математическим представлением, при помощи которого каким-то образом можно объяснить и предсказать поведение электрона. Другие без труда зрительно представляют себе сами волны и поведение электронов в соответствии с тем, как им подсказывают волны при всех видах отражения, рассеяния, взаимодействия и стационарных состояний.

Волновая функция электрона имеет конечную амплитуду во всем доступном для электрона пространстве. Одинокий электрон, блуждающий в космическом пространстве, имеет волновую функцию, распространяющуюся на всю вселенную. С другой стороны, электрон, заключенный внутри атома, обладает волновой функцией, не выходящей за пределы атома. Длина волны λ

волновой функции электрона находится в простой связи с его количеством движения:

$$\lambda = \frac{h}{mv}, \qquad (15\text{-}1)$$

где h — константа Планка ($6{,}6 \cdot 10^{-27}$ *эрг · сек*), выраженная в сантиметрах, если mv выражается в *г · см · сек*. Это соотношение называется уравнением Де-Бройля. Оно утверждает, что медленному электрону соответствует волновая функция, имеющая относительно бóльшую длину волны; быстрый электрон обладает большим количеством движения, большей кинетической энергией и более короткой длиной волны Де-Бройля.

Объемный резонатор возбуждается только на тех частотах, которые соответствуют стоячим волнам внутри него. Подобным же образом потенциальная яма ядра атома удерживает только электроны с такими орбитами, для которых энергия электрона и его количество движения соответствуют длинам волн Де-Бройля, укладывающихся целое число раз на длине орбиты. Это условие осуществляется только для некоторых определенных орбит из многих, которые можно нарисовать в потенциальной яме на рис. 15-2. Таким образом, из бесконечно большого числа электронных орбит, возможных по законам классической механики, разрешаются только немногие. Применив эти идеи на практике, можно вычислить энергию электрона на любой из разрешенных орбит такой простой атомной системы, как водород. Получившийся спектр разрешенных энергетических уровней и энергий показан на рис. 15-4, где уровни названы соответственно их спектроскопическим обозначениям. Следует помнить, что если электрон действительно занимает один из этих энергетических уровней, картина меняется. Если теперь второй электрон подходит к этому нейтральному атому, перед ним открывается совершенно другая энергетическая схема. Глубокой потенциальной ямы не существует для второго электрона за пределами орбиты первого, так как положительный заряд ядра практически нейтрализован отрицательным зарядом первого электрона. Ядра водорода иногда существуют в стабильной связи с двумя электронами одновременно— в виде водородных ионов, но второй электрон очень слабо

удерживается и время жизни такой системы, если только она образуется, обычно мало

На рис. 15-4 электрон изображен занимающим третий снизу энергетический уровень. Эта схема представляет водородный атом в одном из возможных энергетических состояний. Согласно одному из законов природы система, предоставленная сама себе, самопроизвольно стремится перейти в состояние с более низкой энергией. Спустя короткое время электрон, представленный на

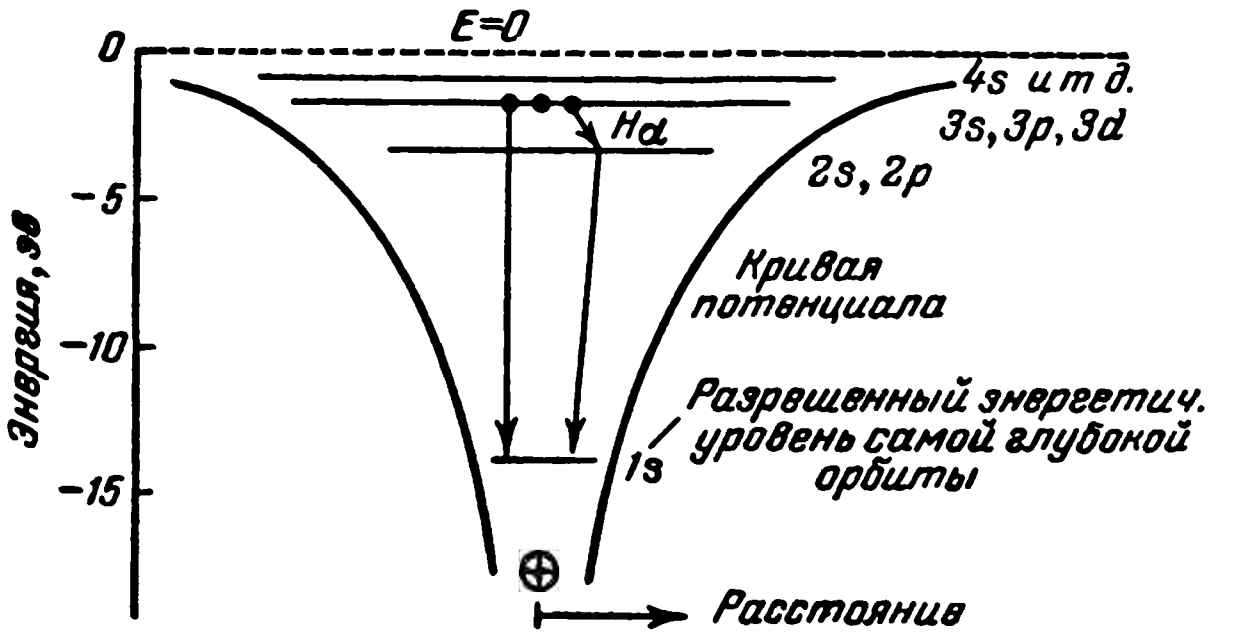


Рис 15-4 На схеме энергетических уровней атома водорода показаны возможные переходы из первоначально возбужденного состояния

энергетической схеме на рис. 15-4, упадет на более низкий уровень и в конечном счете попадет на самый низкий разрешенный уровень, достигнув его либо одним скачком, либо рядом последовательных меньших скачков с задержками на промежуточных уровнях. При каждом таком скачке («переходе» на языке спектроскописта) атом испускает фотон радиации и энергия фотона точно равна разности между энергиями атома в его первоначальном и конечном состояниях. Возможные сопровождаемые излучением фотонов самопроизвольные скачки электрона на расположенные ниже энергетические уровни показаны на рисунке стрелками. В атоме водорода при комнатной температуре электрон почти всегда находится на самом низком энергетическом уровне.

При процессе, обратном только что обсужденному, электрон, первоначально находившийся на самом низком энергетическом уровне, может подняться до одного из более высоких уровней, поглотив энергию фотона из-

лучения, падающего на этом извне. Если падающий фотон обладает энергией, большей, чем вертикальное расстояние от самого низкого энергетического уровня до пунктирной линии $E=0$, он может полностью удалить электрон из области, прилегающей к ядру, которое в результате становится ионом водорода. Освобожденный электрон обладает кинетической энергией, равной разности между положительной энергией, принесенной фо-

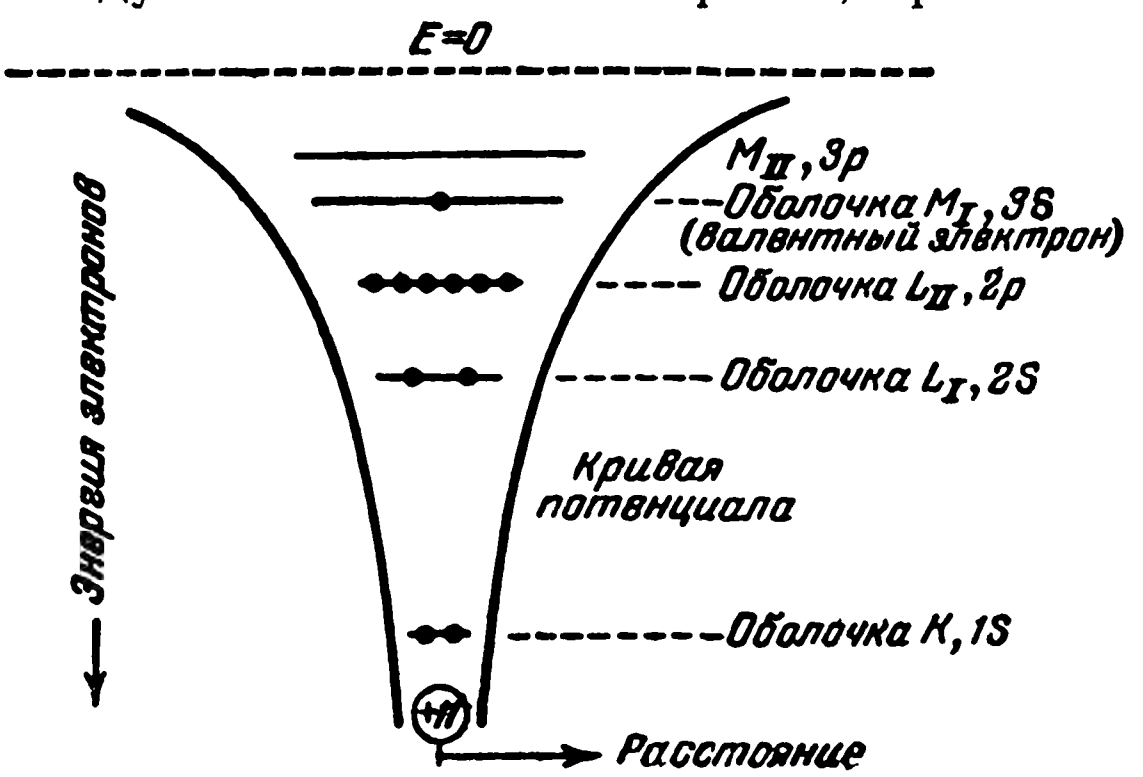


Рис 15-5 Энергетическая схема натриевого атома содержит 11 электронов. Как видно из рисунка, в самом низком энергетическом состоянии атома эти электроны размещаются по различным уровням

тоном, и отрицательной полной энергией уровня, на котором первоначально находился электрон. Электрон затем почти свободно движется в пространстве, пока не встретит частицу с потенциальной ямой, в которую он может упасть.

Рассмотрим теперь энергетическую схему несколько более сложного атома. Ядро атома натрия содержит избыточный положительный заряд, равный 11 электронным зарядам. Соответственно нейтральный атом натрия имеет 11 электронов, распределенных между несколькими разрешенными энергетическими уровнями. Энергетическая схема этого атома может быть изображена, как на рис. 15-5. Если мы возьмем ядро атома натрия и мысленно начнем прибавлять по одному из его 11 электронов, то первые два быстро займут самый низкий энергетический уровень (электронная K-оболочка, если использовать рентгеновское обозначение). При прибавлении третьего электрона вступит в действие еще один закон

квантовой механики, который запрещает более чем двум электронам [1] занимать один и тот же энергетический уровень в одной системе.

Этот закон известен как «принцип исключения» [2]. Третий электрон (а также четвертый) должен поэтому занимать следующий самый низкий энергетический уровень (оболочка L_{I}). Пятый электрон займет третий самый низкий уровень (оболочка L_{II}) и т. д. На рисунке показано, что на этом третьем уровне находятся шесть электронов; это можно рассматривать как совпадение трех энергетических уровней, которые недостаточно разделены у атомов с низким атомным номером, чтобы на рисунке можно было изобразить их раздельными линиями. С добавлением каждого электрона форма потенциальной кривой на рис. 15-5 изменяется, и одновременно изменяется положение всех энергетических уровней, представленных на рисунке, так как добавление нового электрона изменяет суммарный заряд системы. Энергетические уровни изображены так, чтобы показать картину после прибавления 11-го электрона, комплектующего нейтральный атом.

Двенадцатый электрон, движущийся к этому атому извне, совсем «не замечает» потенциальной ямы. Одиннадцать электронов, уже находящиеся в орбитальных оболочках атома, так экранируют заряд ядра, что этот дополнительный электрон не испытывает электрической силы, стремящейся присоединить его к атому. В действительности, если бы он и попытался это сделать, дополнительный заряд, которым он обладает, так исказил бы потенциальное поле атома, что он был бы оттолкнут. Последнее утверждение, справедливое для натрия, не является общим законом для всех атомов. Некоторые электроотрицательные атомы, например атомы кислорода, могут присоединять один или более дополнительных электронов сверх нормального для них комплекта.

На рис. 15-5 допущено некоторое искажение энергетической шкалы, чтобы можно было представить детали схемы, которые при равномерной шкале плохо заметны. Сравнительно с расстоянием между верхними занятыми уровнями расстояние до самого низкого уровня должно

---

[1] Эти два электрона должны иметь противоположные спины.

[2] У нас этот закон часто называется принципом Паули. *Прим ред.*

в действительности измеряться метрами. Представление о положениях различных разрешенных энергетических уровней в разных атомах можно получить экспериментально, наблюдая их оптические и рентгеновские спектры. Найдено, что отрицательная полная энергия электрона на самом низком уровне водородного атома равна приблизительно 13 *эв*; для натрия она равна приблизи-

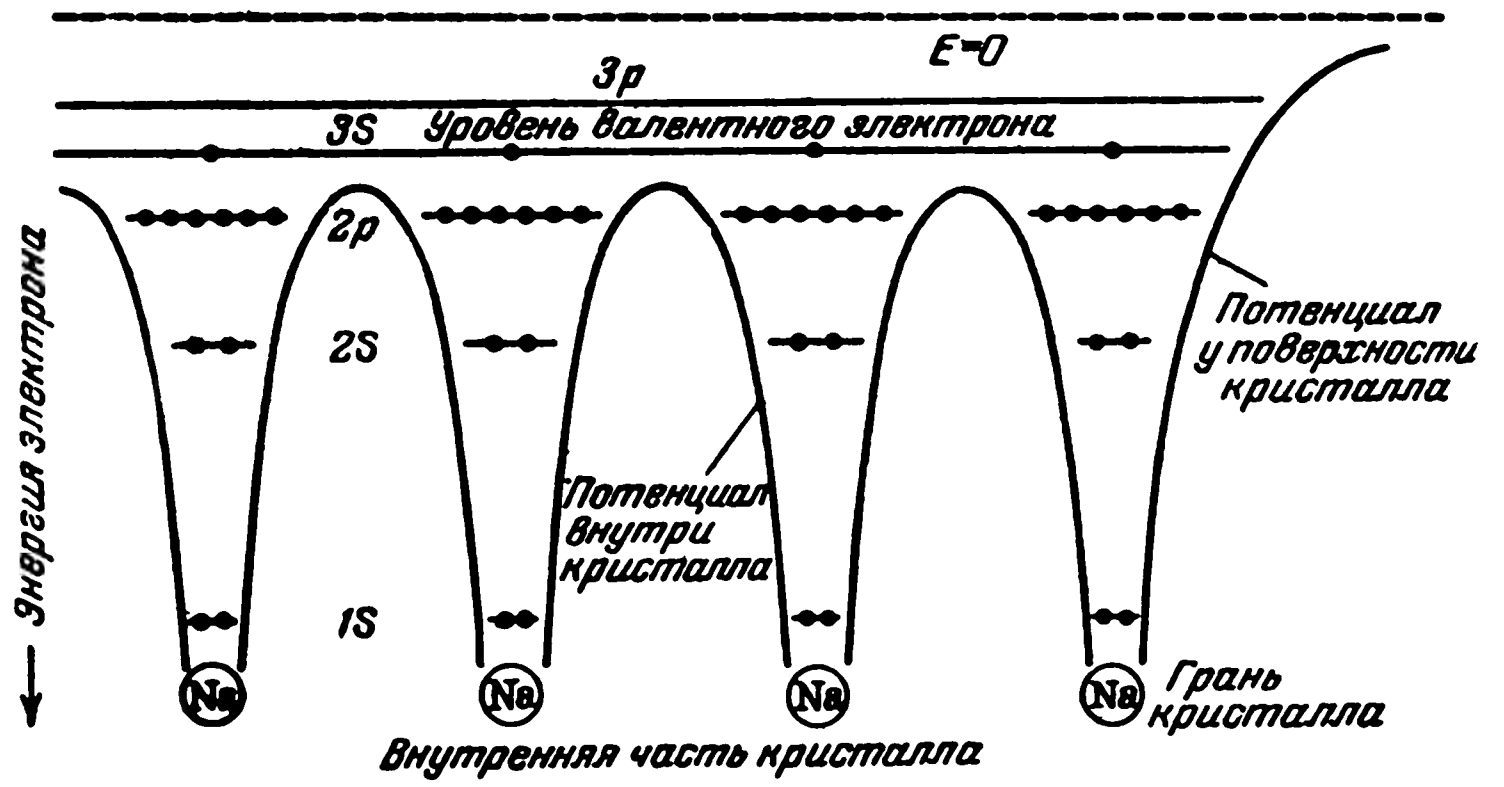


Рис. 15-6. Можно представить себе процессы в кристалле, предположив, что схема энергетических уровней вдоль ряда атомов в куске металлического натрия выглядит так, как показано на чертеже.

тельно 1 100 *эв*; для урана эта величина оказывается порядка 100 000 *эв*.

Заметим, что самый высокий занятый уровень в атоме натрия содержит только один электрон. Это валентный электрон, от поведения которого зависят в большой мере химические и оптические свойства атома. Другие электроны лежат так глубоко в энергетической яме, что они не участвуют ни в обычных химических, ни в термических, ни в оптических процессах, происходящих при изменениях энергии всего лишь на несколько электронвольт.

Предположим, что мы объединили большое число атомов натрия, чтобы образовать небольшой кристаллик. Разобраться в том, что получилось, можно, вычертив кривую потенциальной энергии вдоль цепочки атомов, как это примерно изображено на рис. 15-6. Из-за того, что атомы теперь расположены близко друг к другу, потенциальная энергия в пространстве между

ними не может подняться до $E=0$, как это происходит вне кристалла Потенциальные «бугры» между атомами не поднимаются даже до уровня энергии единственного валентного электрона в каждом атоме. Поэтому ничто не препятствует валентному электрону, первоначально принадлежавшему данному атому, покинуть свой материнский атом и начать свободно блуждать по кристаллу. В частности, если внутри кристалла создать электрическое поле, эти свободные электроны должны увлекаться в одном направлении этим полем и их направленное движение должно привести к возникновению электрического тока. Единственное требование к возможности подобной свободы валентных электронов заключается в том, чтобы ни в одном месте кристалла не скапливалось намного больше или меньше электронов, чем это соответствует приблизительно местной нейтральности результирующего заряда. Можно думать, что мы создали таким образом с помощью основных представлений физики простую картину того, откуда появляются свободные электроны, ответственные за электропроводность таких типичных металлов, как натрий.

Рис. 15-6, каким бы логичным он ни казался, в какой-то степени является чрезмерно упрощенным, и действительное поведение электронов внутри кристалла требует дальнейшего рассмотрения. Заметим, что все валентные электроны, различных атомов одного и того же элемента обладают, как показано на рис. 15-6, одинаковой энергией. Но так как они больше не принадлежат ни одному определенному атому. то их надо считать принадлежащими всему кристаллу, а кристалл рассматривать как единую систему. Однако считать, что все эти электроны имеют одинаковую энергию в одной системе, было бы нарушением принципа исключения.

В действительности при объединении атомов и образовании кристалла происходит следующее: дискретные энергетические уровни, обычно занимаемые электронами в отдельных атомах, расщепляются на энергетические зоны, как показано на рис. 15-7. Каждая зона состоит из большого числа очень близко расположенных дискретных энергетических уровней. В каждой зоне столько таких уровней, сколько атомов в кристаллическом образце, а в кристалле столько таких зон, сколько энергетических уровней в изолированном атоме этого ве-

щества. Трудности, связанные с принципом исключения, этим устраняются, так как различные валентные электроны атомов теперь могут занять различные энергетические уровни в пределах зоны 3*s*. В случае натрия ширина вышележащих зон такова, что зоны 3*s* и 3*p* перекрывают друг друга. В кристалле, содержащем N одновалентных атомов, содержится N валентных электронов. Валентная зона в кристалле в этом случае имеет N близко расположенных энергетических уровней и может быть заполнена $2N$ электронами, так как каждый уровень может принять два электрона. Уровни этой зоны должны тогда быть заполнены лишь наполовину, так как в изолированных атомах натрия имеется лишь по одному валентному электрону на один валентный уровень. В действительности у натрия «дно» перекрывающей 3*p* зоны соответствует несколько меньшей энергии, чем максимальная энергия электронов в наполовину заполненной зоне 3*s*. В результате некоторые из 3*s* электронов переходят на нижние уровни зоны 3*p* и, начиная с этого уровня, обе зоны заполняются вверх одновременно, пока не исчерпываются все имеющиеся электроны.

В изолированном атоме существует ряд энергетических уровней, расположенных выше тех, которые показаны на рис. 15-5. Это приводит к появлению дополнительных расположенных выше перекрывающихся зон, которые на рис. 15-7 не показаны, чтобы изображение было по возможности простым.

Ширина энергетической зоны между ее верхним и нижним уровнями определяется формой кристалла и материалом, из которого он состоит. Валентные зоны большинства твердых веществ имеют ширину всего лишь в несколько электронвольт. Ширина энергетической зоны не зависит от размера образца или общего числа атомов, заключенных в нем. В двух образцах кристалла одного и того же вещества, из которых один в 2 раза больше другого, первый имеет в 2 раза больше энергетических уровней в каждой зоне, чем второй, и эти уровни расположны в 2 раза теснее друг к другу в зонах и вмещают в 2 раза больше электронов, но соответствующие края верхней и нижней зон имеют одинаковое взаимное расположение по шкале энергий в обоих образцах.

Представлялось бы весьма искусственным утверждением, что когда атомы собираются вместе, образуя кристалл, дискретные энергетические уровни в отдельных атомах расщепляются на зоны уровней просто для того, чтобы избежать нарушения принципа исключения. Полное объяснение этого расщепления можно найти в квантовой физике, но качественное представление о физике этого явления можно получить, рассмотрев некоторую механическую аналогию. Представим себе ряд из N оди-

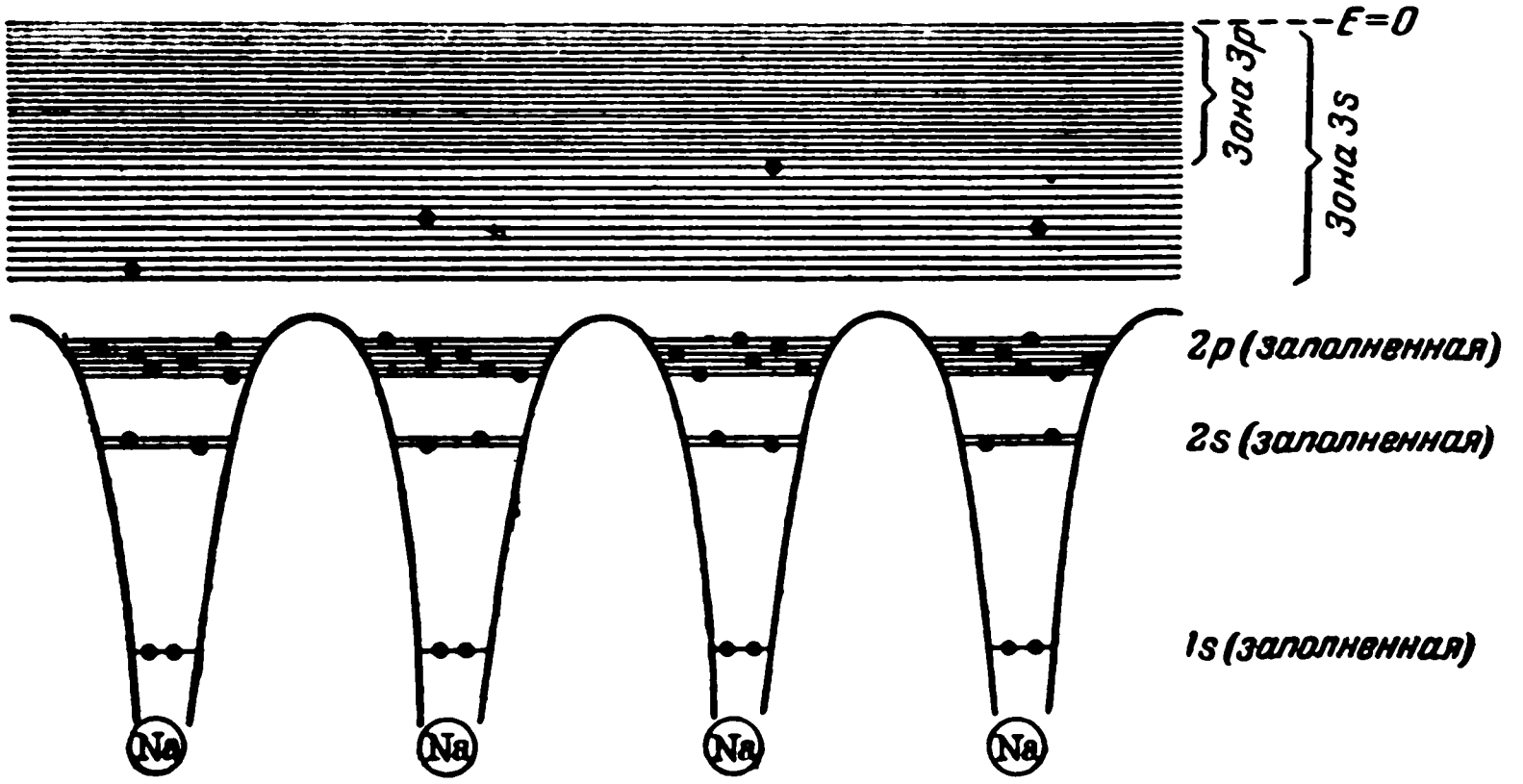


Рис 15-7. В твердом металлическом натрии энергетические уровни атомов распадаются на зоны уровней.

наковых механических генераторов, действующих независимо друг от друга и создающих колебания с одной и той же частотой. Теперь предположим, что между каждым генератором и его соседями в ряду создана механическая связь. Тогда окажется, что ранее независимые генераторы колебаний образуют одну систему, частота колебаний которой может иметь одно из N различных значений в спектре N близко расположенных друг к другу частот, сгруппированных вокруг первоначальной одной частоты изолированного генератора колебаний. Чем сильнее связь между генераторами, тем шире образующаяся зона и тем дальше отстоят друг от друга в зоне составляющие ее N частот.

Подобным же образом в кристалле между электронами некоторого атома и электронами его соседей существует волно-механическое взаимодействие. Эта

связь, естественно, сильнее для электронов на выше расположенных уровнях. Эти уровни соответствуют самым внешним электронным оболочкам атомов, и их волновые функции постоянно взаимодействуют с волновыми функциями соответствующих уровней в соседних атомах. Следовательно, расщепление энергетических уровней в этих верхних зонах кристалла самое сильное, а ширина зон оказывается наибольшей. Вследствие существования потенциальных барьеров между атомами в кристалле волно-механическое взаимодействие между

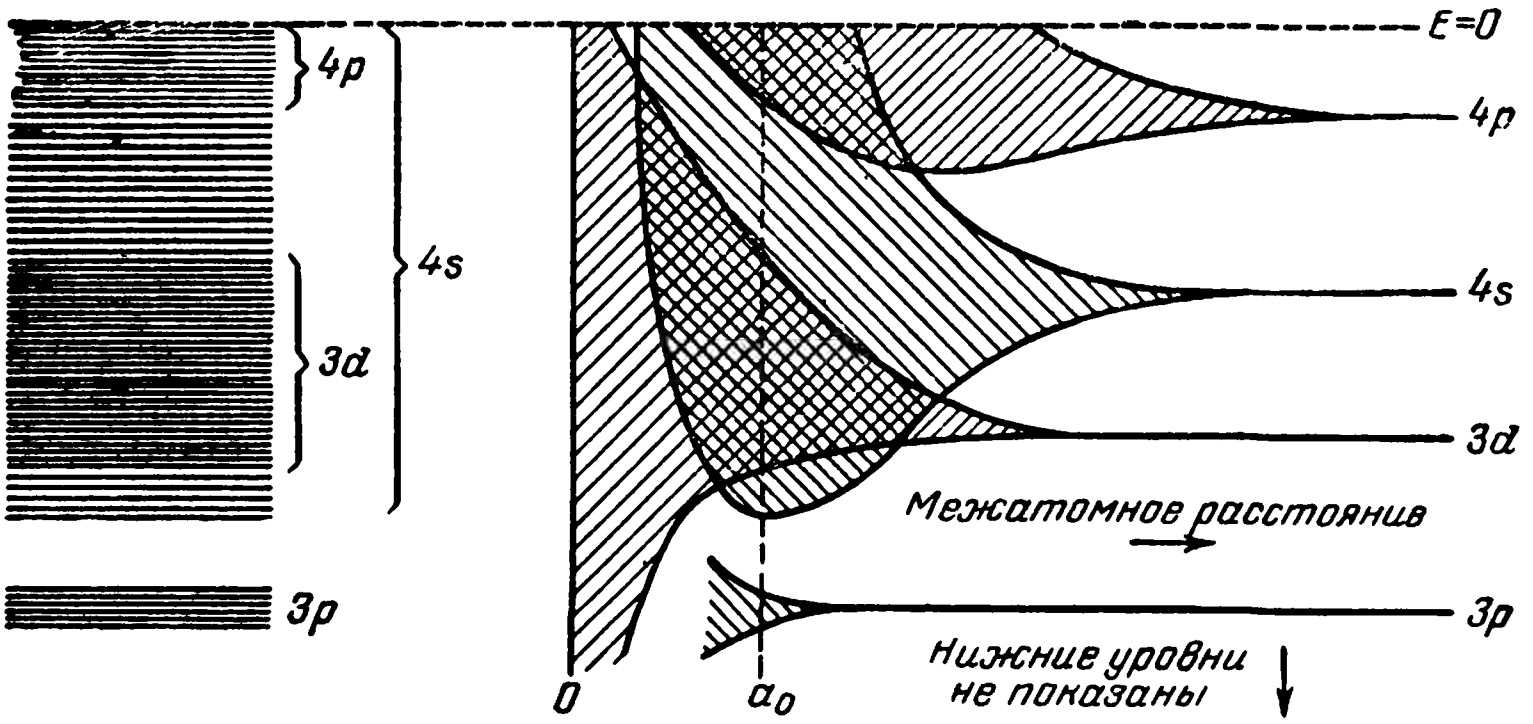


Рис. 15-8. В металлической меди (слева) верхние энергетические зоны перекрываются довольно сложным образом Расщепление энергетических уровней атомов меди (справа) зависит от межатомного расстояния.

электронами ниже лежащих уровней (внутренних орбит) соседних атомов значительно слабее и нижние зоны у́же. Для самых низких зон (для самых внутренних электронных оболочек в атомах) взаимодействие почти не существует, так как эти электроны защищены от внешнего мира внешними электронными оболочками. Следовательно, самые низкие зоны очень узки — едва ли заметно шире дискретных уровней в изолированных атомах. На рис. 15-7 сделана попытка показать качественно сравнительную ширину этих различных энергетических зон.

На рис. 15-8 слева показана схема валентных уровней для более сложного вещества — металлической меди, а справа — как возникают зоны из энергетических уровней изолированных атомов меди. На правом краю

этой диаграммы энергетические уровни дискретны и узки, как у отдельных изолированных атомов меди. По мере сближения атомов верхние энергетические уровни начинают взаимодействовать и расщепляются, образуя зоны. При большем сближении атомов усиливается расщепление следующих энергетических уровней и т. д. К тому моменту, когда атомы уплотняются до межатомного расстояния a_0, соответствующего нормальному кристаллу меди, зоны $3d$, $4s$ и $4p$ настолько расширяются, что они начинают перекрываться, как показано на рисунке, в то время как низкие энергетические зоны расширяются в меньшей степени. Чтобы вызвать заметное расщепление самых внутренних энергетических уровней оболочки K, надо было бы применить физически неосуществимое давление, чтобы заставить атомы подойти друг к другу намного ближе, чем они находятся обычно в кристалле меди.

В изолированном атоме меди все энергетические уровни заполнены до уровня $3d$ включительно. На уровне $4s$, который мог бы вместить два электрона, имеется только один валентный электрон на атом. В кристаллической меди соответственно заполнены все зоны до $3d$ включительно. Зона $4s$ заполнена до уровня, расположенного немного выше уровня ее перекрытия с верхним краем зоны $3d$.

На рис. 15-9 показана зонная схема ионного кристалла соединения закиси меди Cu_2O. В этом соединении атомы меди так далеки друг от друга, что зоны $3d$ и $4s$ не перекрываются. Кислородная зона $2p$ лежит, как показано, ниже зоны $3d$ меди и содержит шесть электронов на один атом кислорода. Электронная структура кислорода обычно в этой зоне содержит только четыре электрона на атом. Однако если каждый из двух атомов меди отдает свой $4s$-й электрон в зону $2p$ каждого атома кислорода, то в результате образуется структура, изображенная на рис. 15-8, в которой кислородная зона $2p$ совершенно заполнена, а зона $4s$ меди совершенно пуста. При этом оказывается, что в случаях металлической меди и натрия есть пустые электронные уровни сразу же над верхним из заполненных электронами уровней, тогда как в закиси меди существует энергетический зазор между вершиной заполненной зоны $3d$ и дном пустой зоны $4s$.

Как мы увидим в следующей главе, различие между металлами и неметаллами может быть объяснено расположением и степенью заполнения энергетических зон в твердом веществе. Картина, представленная на

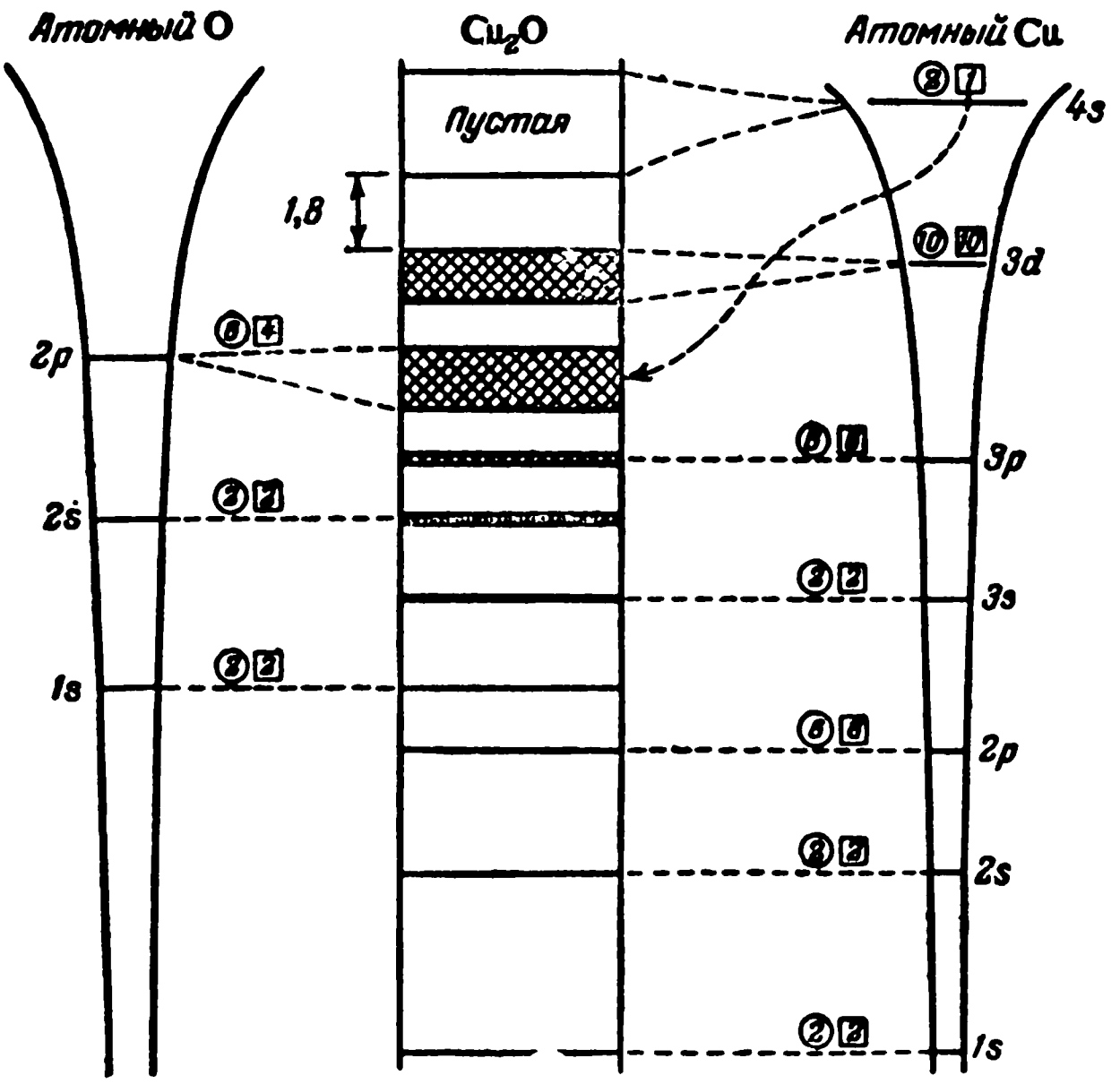


Рис 15-9 Энергетические зоны закиси меди возникли из энергетических уровней меди и кислорода Все зоны Cu_2O полностью заполняются до зоны 3*d* меди Электроны с уровней 4*s* меди перешли в зону 2*p* кислорода, остающуюся незаполненной Числа в кружках обозначают число электронов на атом, которое может вместить каждая зона Числа в прямоугольниках указывают действительные числа электронов на атом в зонах.

рис. 15-9, является картиной стехиометрически чистой закиси меди. Наличие примесей, отклонения от стехиометрического состава или же дефекты в решетке усложняют эту простую картину, как будет выяснено в дальнейшем.

Мы привели достаточно примеров, чтобы иллюстрировать концепцию энергетических зон с атомистической точки зрения.

Теперь обратимся к построению картины энергетических зон из рассмотрения электронного коллектива. Эти два метода приближения даны здесь потому, что в каждом из них есть свои достоинства.

15-2. ЗОННАЯ ТЕОРИЯ С ТОЧКИ ЗРЕНИЯ КОЛЛЕКТИВНОГО ЭЛЕКТРОННОГО ПРИБЛИЖЕНИЯ

В § 15-1 картина энергетических зон была построена с помощью атомистической теории внимание концентрировалось на энергетических уровнях отдельных атомов и изменениях уровней, происходящих в результате группирования большого числа атомов при образовании

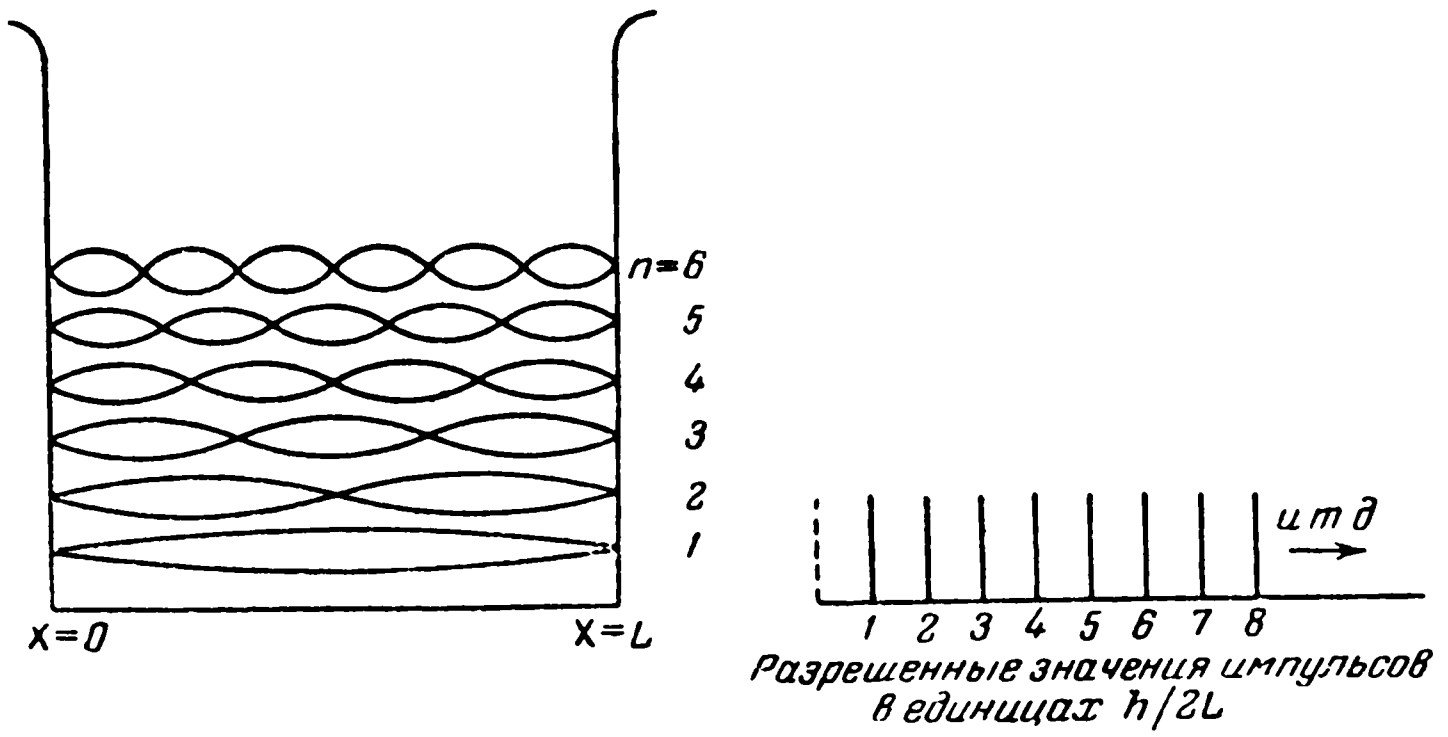


Рис. 15-10 Внутри ящика (слева) могут поддерживаться только стоячие электронные волны, что приводит к спектру разрешенных значений импульсов электронов, показанному справа.

твердого кристалла. Можно подойти к решению этого вопроса и с других позиций, рассматривая поведение газа свободных валентных электронов внутри области, ограниченной поверхностью кристалла, и возмущение их поведения, вызываемое присутствием атомных ядер и периодичностью кристаллической решетки.

Приближение, исходящее из теории коллективизированных электронов, применимо только для валентных электронов атомов кристалла, так как на расположенных ниже уровнях эти электроны связаны со своими материнскими атомами и не участвуют в электронном газе.

Рассмотрим одномерный потенциальный ящик, имеющий вертикальные стенки и плоское дно, как показано на рис. 15-10 слева. Этот ящик представляет собой силь-

но идеализированный аналог воронкообразной потенциальной ямы отдельного атома, расширенной до такой степени, чтобы она могла представлять внутреннюю часть всего одномерного образца кристалла. Вертикальные стенки по сторонам ящика изображают определяющий работу выхода скачок потенциала на поверхностях кристалла, удерживающий электроны внутри.

Представим себе, что этот ящик содержит свободные валентные электроны и достаточно равномерно распределенный положительный заряд, обеспечивающий нейтральность пространства. Этот положительный заряд мы будем считать не имеющим структуры, а движение электронов в ящике — происходящим без вязкостного сопротивления.

Внутри этого ящика могут существовать лишь такие электронные волновые функции, половины длин волн которых укладываются на длине L ящика целое число раз, т. е. только волновые функции, соответствующие стоячим волнам внутри потенциального ящика. Несколько волновых функций, соответствующих самым низким разрешенным значениям импульсов в ящике, изображены на рис. 15-10, а справа представлена диаграмма спектра разрешенных значений импульсов электронов.

Длины стоячих волн, могущих существовать в этом ящике, выражаются следующим образом:

$$\lambda_n = \frac{h}{mv_n} = \frac{2L}{n}\,;\ n = 1, 2, 3 \text{ и т. д.} \qquad (15\text{-}2)$$

Соответствующие разрешенные импульсы электронов равны:

$$mv_n = \frac{nh}{2L}\,;\ n = 1, 2, 3 \text{ и т. д.,} \qquad (15\text{-}3)$$

а соответствующие кинетические энергии

$$KE = \frac{1}{2}\,mv_n^2 = \frac{(mv_n)^2}{2m} = \frac{n^2h^2}{8mL^2}\,;\ n = 1, 2, 3 \text{ и т. д.} \qquad (15\text{-}4)$$

Полезно рассмотреть график кинетической энергии в функции импульса для разрешенных импульсов электронов в потенциальном ящике. Такой график показан на рис. 15-11, где каждый разрешенный электронный уровень соответствует двум точкам на кривой, причем

эти две точки обладают импульсами, равными по величине, но противоположными по направлению. Из-за квадратичной зависимости энергии от импульса эти точки лежат на параболе.

Электроны внутри потенциального ящика занимают самые нижние разрешенные уровни на параболе и по ее сторонам, пока не будут исчерпаны все электроны Только два электрона могут занять один уровень; эти два электрона должны иметь противоположные спины. На

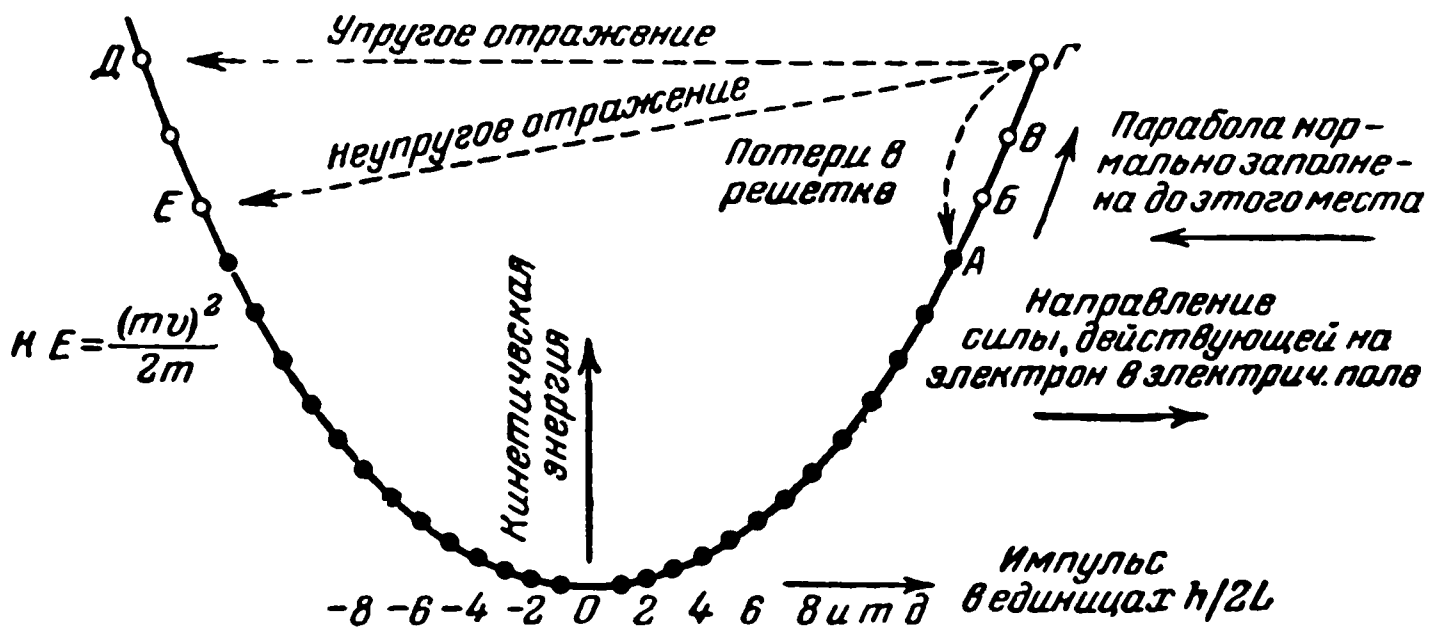


Рис 15-11 Каждая разрешенная электронная волновая функция на рис 15-10 дает две точки на параболической кривой в пространстве энергий-импульсов, эта пара точек изображает равные по величине, но противоположно направленные импульсы

рис. 15-11 заполненные уровни обозначены сплошными точками; кружочками отмечены незанятые уровни, расположенные выше по сторонам параболы.

Простая модель проводимости металла прямо следует из этой картины. Заметим, что при отсутствии какого-либо возмущающего воздействия обе ветви параболы заполняются одинаково. Направо и налево движется одинаковое количество электронов с одинаковыми скоростями, так что распределение по импульсам валентных электронов, несмотря на хаотический беспорядок их движений, находится все время в динамическом равновесии и никакого результирующего тока не возникает. Если же теперь наложить электрическое поле, побуждающее электроны двигаться вправо (рис. 15-11), то электрон, находящийся на вершине распределения в точке *А*, может ускоряться полем, совершая последовательные скачки на уровни со все бóльшими импульсами: *Б*, *В*, *Г*. Это ускорение электрона сообщает ему

большую скорость в направлении поля, что нарушает симметрию заполнения электронами уровней на параболе и приводит к появлению в образце результирующего электрического тока.

Если в результате ускорения электрон переведен в более высокое энергетическое состояние, например *Г*, то это может привести к ряду последствий. Он может продолжать ускоряться, переходя на все более высокие уровни параболы. Или же может произойти неупругое столкновение между электроном и одной из гармоник тепловых колебаний кристаллической решетки и он может попасть обратно на вершину распределения возле точки *А*, теряя приобретенную в электрическом поле энергию и нагревая кристалл отдачей ему этой энергии. Этот процесс внешне выражается в выделении джоулева тепла. Электрон может также испытать полностью упругое взаимодействие с решеткой и рассеяться (отразиться) так, что при изменении направления импульса на обратное он внезапно обнаруживается на диаграмме в точке *Д*. Наконец, он может неупруго отразиться и оказаться на диаграмме в точке *Е*, потеряв при отражении некоторую энергию. Пока поле включено, распределение электронов на параболе смещено вправо, причем это смещение поддерживается на небольшой установившейся величине за счет трех только что описанных восстановительных процессов. Это смещение соответствует постоянному электрическому полю, направленному вправо.

Когда электрон, соответствующий вершине распределения на параболе в точке *А*, ускоряется полем в направлении *Б*, *В*, *Г*, то в точке *А* остается незанятое состояние, в которое при ускорении может перейти близлежащий расположенный ниже на параболе электрон, оставляя на своем месте новое незанятое состояние, в которое может перейти следующий близлежащий электрон, и т. д Это незанятое состояние становится положительной дыркой, стремящейся перемещаться вниз и влево по параболе, в то время как электроны движутся вверх и вправо. Эта дырка обычно не перемещается сколько-нибудь далеко в этом движении, а лишь до тех пор, пока один из упомянутых выше восстановительных механизмов не придет в действие, стремясь восстановить симметрию в расположении электронов на параболе.

Согласно нарисованной картине ускорение электрона происходит не непрерывно, как в случае совершенно свободной частицы, а скорее в виде последовательности дискретных скачков ко все более высоким скоростям в направлении ускорения под действием поля. Может показаться, что это идет в разрез с нашими интуитивными предположениями относительно того, что может произойти, но природа атомных и электронных явлений полна неожиданностей. В трехмерном кристалле соседние энергетические состояния отстоят друг от друга примерно на 10^{-22} *эв*, так что дискретные скачки оказы-

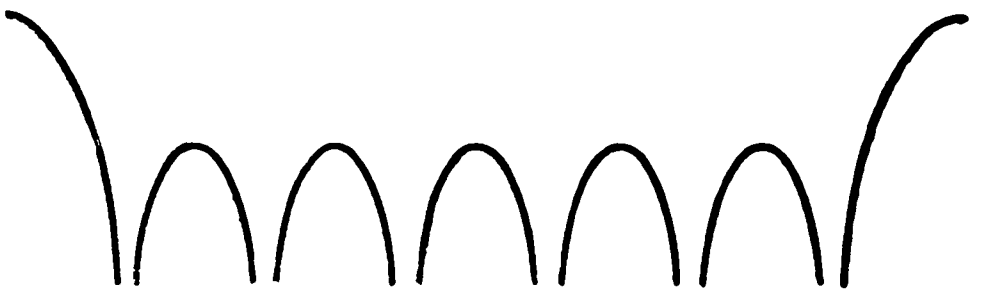


Рис. 15-12. Твердый кристалл является потенциальным ящиком, дно которого имеет периодическую структуру.

ваются очень небольшими и ускорение частицы может рассматриваться как квазинепрерывное.

До сих пор мы говорили об электронном газе, заключенном в плоскодонном потенциальном ящике. Примем теперь во внимание периодичность потенциала дна ящика, вызванную правильным расположением атомных ядер в одномерном кристалле. Внутренность ящика в этом случае выглядит так, как на рис. 15-12. Периодический потенциал дна ящика возмущает волновые функции электронов, которые больше не имеют характера стоячих синусоидальных волн, а являются стоячими волнами более сложных гармонических форм. В результате этого усложнения происходит такое возмущение поведения электронов, что график зависимости кинетической энергии его от импульса для разрешенных уровней электронов имеет разрыв при некоторых величинах импульсов. На рис. 15-13 показан такой график зависимости возмущенной энергии от импульса. Видно, что график в основном соответствует параболе, но отличается от нее вблизи разрыва. Разрешенные электронные уровни обозначены на рисунке сплошными точками и кружками. На этом рисунке, однако, допущены вольности в отношении количества уровней. Для одномерного

кристалла протяженностью 1 *см* на каждом сегменте кривой должно быть от 10^7 до 10^8 точек — одна для каждого атома в образце. Увеличение образца вдвое дало бы нам в 2 раза большее число разрешенных уровней: точки на этом рисунке были бы вдвое ближе друг к другу и их было бы в 2 раза больше. Однако *величины им-*

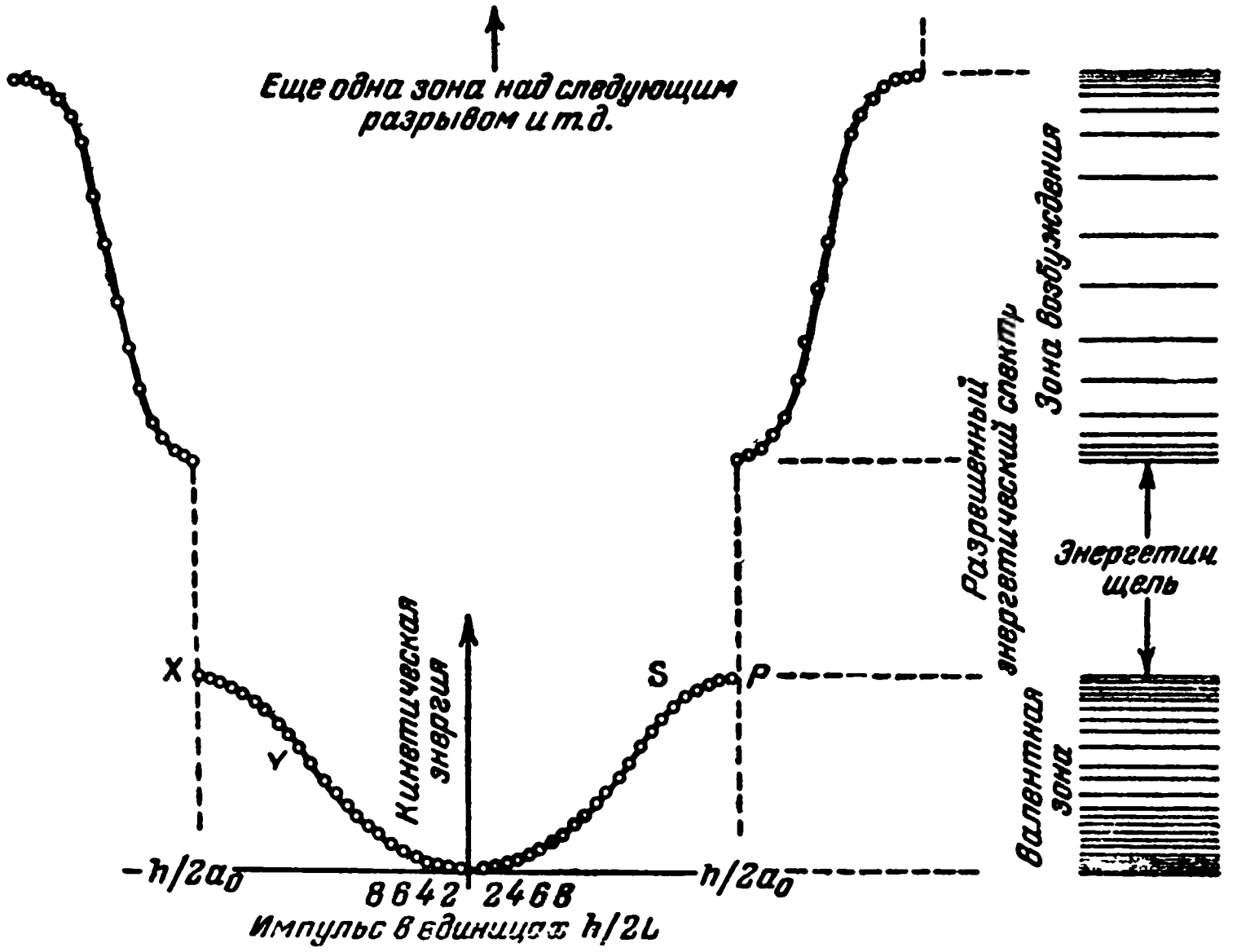


Рис 15 `3 В твердом кристалле атомная периодичность потенциала создает на параболе энергии-импульса показанные здесь искажения

пульсов, соответствующих разрывам, и величины энергии на верхних и нижних краях разрешенных энергетических зон остаются неизменными независимо от числа атомов в образце, если с изменением числа атомов межатомное расстояние не изменяется.

Сплошные участки кривой, заключенные между разрывами, образуют зоны разрешенных энергетических состояний, отделенные энергетическими щелями (запрещенными зонами), образующимися в разрывах. Схема разрешенных энергетических уровней для данного случая представлена справа на рис. 15-13. Здесь показана качественно картина энергетических зон, к которой

в предыдущем параграфе уже привела нас атомистическая теория. Приведшая нас к рис. 15-13 модель, основанная на приближении коллективизированных электронов, была построена в одном измерении. Обобщив эту модель на случай трех измерений, мы получим более точную картину, приближающуюся в бóльших деталях к результатам, полученным из атомистического подхода *

Величины импульсов, при которых имеют место разрывы на рис. 15-13, — это те, для которых половина длины волны волновых функций Де-Бройля укладывается целое число раз в постоянной кристаллической решетки a_0 Самый низкий электронный уровень у вершины параболы соответствует волновой функции, образующей стоячую волну, в которой одна полуволна укладывается на протяжении целого кристалла. Электронный уровень у первого разрыва соответствует электрону с таким импульсом, что его волновая функция образует стоячую волну, имеющую узлы в каждой ячейке решетки. Электрон, соответствующий этой волновой функции, имел бы импульс

$$mv = \frac{h}{2a_0} . \qquad (15\text{-}5)$$

Такая волновая функция, однако, не может распространяться по решетке. Волна не распространяется в периодической структуре, имеющей протяженность периода структуры, равную половине длины волны. Если этот процесс не сопровождается потерей энергии, то волна полностью отразится, пройдя расстояние, равное

---

* Последние данные [Л 2] указывают на несколько бóльшую сложность явлений, чем в этой простой картине Ширина и положение энергетических зон зависят от кристаллографического направления, для которого ведется анализ Особенно существенно это для пространственно анизотропных кристаллов, так как в этих кристаллах межатомные расстояния и распределение потенциала между атомами зависят от направления в кристалле Однако считается, что даже для кремния и германия, которые обычно рассматриваются как вещества изотропные, вдоль определенных кристаллографических направлений кривая зависимости энергии от импульса вместо того, чтобы быть параболой, имеет минимумы при конечных положительных и отрицательных значениях импульса с пологим энергетическим горбом при $mv=0$ В нашей элементарной схеме мы опустим рассмотрение этих тонкостей и будем рассматривать лишь основные особенности энергетической схемы.

нескольким периодам структуры. Это отражение волн, на которое периодическая структура настроена, лежит в основе брэгговского отражения рентгеновых лучей. Оно также является причиной возникновения зон непропускания в цепочке из одинаковых резонансных фильтров. Отражение решеткой волновых функций, имеющих точно указанное отношение длины волны к постоянной решетки и сильное взаимодействие решетки с волновыми функциями, имеющими очень близкое к указанному отношение этих величин, и является причиной возмущений и разрывов в кривых зависимости энергии от импульса на рис. 15-13. С помощью графика энергии в функции импульса могут быть выявлены еще некоторые интересные особенности поведения электронов в зонах. Так как

$$KE = \frac{1}{2} mv^2 = \frac{(mv)^2}{2m}, \qquad (15\text{-}6а)$$

то

$$\frac{d\,(KE)}{d\,(mv)} = \frac{mv}{m} = v - \text{ крутизна кривой;} \qquad (15\text{-}6б)$$

$$\frac{d^2\,(KE)}{d\,(mv)^2} = \frac{1}{m} - \text{кривизна кривой.} \qquad (15\text{-}6в)$$

Из этого аналитического определения следует, что крутизна кривой зависимости энергии от импульса дает скорость электрона, занимающего разрешенный уровень в месте, где эта крутизна рассматривается, и что крутизна кривой в этой точке обратна величине массы электрона. Так как эта кривая отличается от параболической, характерной для классического движения электронов в пространстве с непериодическим потенциалом, то очевидно, что вблизи разрывов на рис. 15-13 имеет место особое поведение электронов в отношении их скоростей и масс. Рассмотрим, например, электрон на одном из уровней вблизи точки S на этом рисунке. Пусть импульс этого электрона увеличивается так, что он перемещается на уровень, расположенный правее исходного. Здесь его скорость меньше, чем она была до этого уровня [уравнение (15-6б)], так как крутизна кривой E в зависимости от mv уменьшается по мере приближения к разрыву. Таким образом, движение электрона замедляется при увеличении его импульса.

Можно до некоторой степени понять причину этого явно противоречащего нашим знаниям поведения, если учесть следующее: кристаллическая решетка действует на электронные волны во многом подобно фильтру, пропускающему низкие частоты. Если электрон находится в одном из состояний с импульсом, близким к точке S (рис. 15-13), то приложенное постоянное электрическое поле стремится увеличить его импульс в направлении действующей в поле силы и длина волны Де-Бройля становится короче. Так же как в случае движения волны в электрическом полосовом фильтре, скорость распространения электронной волны уменьшается по мере того как длина ее волны уменьшается, приближаясь к длине волны отсечки кристалла. Поэтому движение электрона замедляется по мере роста его импульса от точки S к первому разрыву на графике зависимости E от mv. Это может произойти, однако, только если эффективная масса электрона отрицательна. Отрицательное значение кривизны кривой в этой области [(уравнение (15-6в)] совпадает с отрицательным значением электронной массы, соответствующим этому участку кривой.

Этот результат чрезвычайно поразителен, так как одно из основных понятий в курсах физики — это нерушимость закона сохранения массы. Электрон мог бы вести себя так, как описывалось выше, еще и в том случае, если бы его заряд мог изменять знак, но закон сохранения заряда всегда представлялся тоже нерушимым. Эту теоретическую трудность можно частично разрешить, сказав, что «масса», о которой идет речь, необязательно совпадает с результатом, который мы получим, если сможем как-то, не нарушая движения электронов в твердом веществе, привести его в равновесие и взвесить. Конечно, массу следует рассматривать здесь как отношение силы поля, действующей на электрон, к испытываемому им ускорению. Если в результате чего-то это отношение оказывается непостоянным, то мы объясняем это поведение так, как если бы масса электрона была переменной. В этом смысле можно считать что электрон у верхнего края зоны *ведет себя так, как если бы* у него была отрицательная масса, как это вытекает из уравнения (15-6б) и (15-6в).

Эта концепция «эффективной» массы представляется

интересной и смелой. Эффективная масса электрона зависит от того, на каком уровне электрон появляется в зоне, и от влияния, которое периодический потенциал решетки оказывает на кривизну кривой E в зависимости от mv в рассматриваемой точке. Эффективные массы могут быть любого знака и могут значительно отличаться от массы покоя электрона в свободном пространстве.

Чтобы получить лучшее ощущение характера движения электрона в твердом теле и его странного и непонятного поведения, рассмотрим уединенный электрон в свободной зоне, начинающий движение из состояния покоя в нижней части кривой на рис. 15-13 и увлекаемый электрическим полем вправо. Предположим далее, что в кристаллической решетке электрон не теряет энергию, т. е. ничто не препятствует движению электрона в поле. Посмотрим, что произойдет со скоростью электрона и его эффективной массой в результате непрерывного воздействия поля. Мы можем проследить изменение этих параметров с помощью кривых на рис. 15-14, где энергия, скорость и масса показаны графически на основе уравнений (15-6а), (15-6б) и (15-6в) в функции импульса. Когда электрон начинает двигаться, его эффективная масса и ускорение почти постоянны соответственно линейному возрастанию скорости с ростом импульса (если не считаться с разрывами кривой, связанными с последовательными скачками от одного состояния импульса к другому с низких уровней на все более высокие). Это «нормальное» поведение продолжается до тех пор, пока электрон не достигнет уровня с таким импульсом, где взаимодействие электронной волны и периода решетки начинает сказываться на величине скорости волны. Когда электрон приближается к точке S на энергетическом графике, кривая отклоняется от параболической, ускорение электрона падает и его эффективная масса возрастает. В точке перегиба S эффективная масса вдруг изменяется от очень большого положительного значения до очень большого отрицательного значения, и движение электрона начинает замедляться, хотя его импульс возрастает с течением времени с постоянной быстротой все время, пока поле действует. Это замедление показано ходом кривой скорости, по мере того как импульс электрона изменяется от точки S до точки P. Когда электрон приходит в точ-

ку P, его скорость делается равной нулю и он останавливается на своем пути, несмотря на то, что его импульс в этот момент наибольший. В этой точке электрон испытывает брэгговское отражение в решетке и появляется на параболе в точке X, где его скорость все еще равна нулю; его эффективная масса все еще отрицательна, а вектор его импульса изменен отражением на обратный. Теперь электрон начинает ускоряться *против* силы, действующей на него в поле, в результате его отрицательной эффективной массы. Когда он доходит до точки Y, его масса снова становится положительной и направленная обратно скорость уменьшается до тех пор, пока его импульс не пройдет через нуль в нижней части кривой, где он снова оказывается в покое и готов повторить описанный выше процесс, если электрическое поле продолжает на него действовать. Электрон, таким образом, почти все время находится в движении, но никогда никуда не приходит подобно абсолютно упругому мячу,

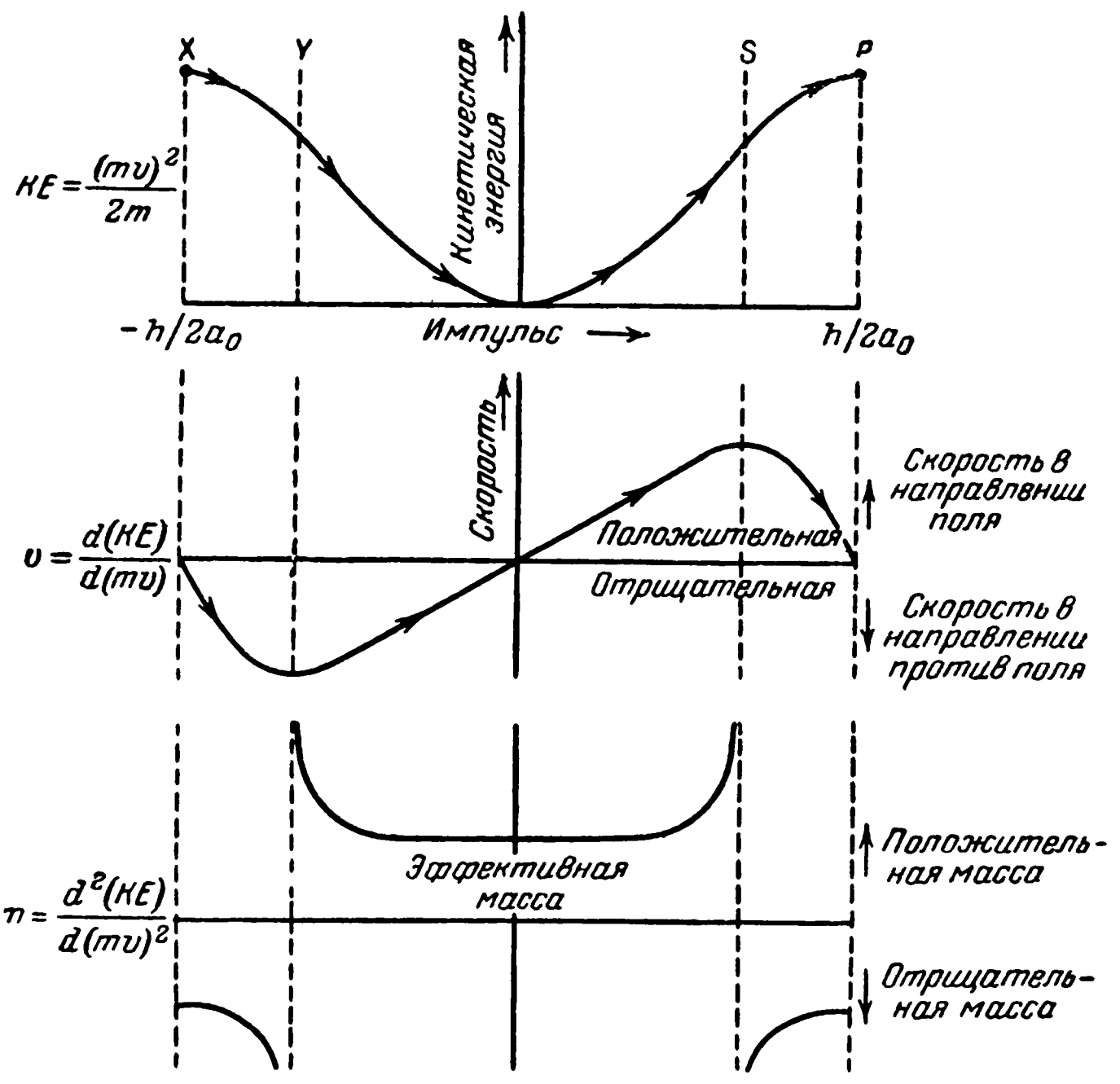


Рис. 15-14 Энергия, скорость и эффективная масса электрона в зависимости от импульса в энергетической зоне твердого тела

прыгающему на абсолютно упругом полу, пока не будет уничтожена сила тяжести. Если не помешать этому периодическому движению, например потерей энергии при столкновениях, электрон не будет создавать постоянного электрического тока в этом веществе.

В энергетической зоне, *в которой все уровни полностью заняты,* нет свободных уровней, на которые электроны могли бы быть переведены в результате какого-нибудь рассеивающего энергию механизма, о котором говорилось раньше; в этом случае все электроны находятся в таком циклическом движении все время, пока приложено электрическое поле, не создавая, однако, никакого результирующего тока, несмотря на то, что эти электроны «свободны», т. е. не связаны со своими материнскими атомами. И только когда зона частично заполнена или частично пуста, подобное взаимодействие меняет их поведение, приводя к асимметрии в движениях электронов в этой зоне, что ведет к появлению результирующего электрического тока в образце. Созданная выше картина проводимости металла предполагает существование зоны, лишь частично заполненной электронами, в которой над занятыми уровнями находятся свободные. Именно за счет таких обычно незанятых уровней действует механизм проводимости.

Так как масса электронов вблизи вершины полностью заполненной энергетической зоны ведет себя, как группа частиц, обладающих отрицательной массой и отрицательным зарядом, удаление одной из этих частиц оставляет дырку, имеющую эффективную положительную массу и положительный заряд. Это новое раскрытие природы положительной дырки проливает дополнительный свет на ситуацию, складывающуюся при создании зонной картины с позиций атомистической теории.

В модели, которую мы создали, основываясь на теории коллективизированных электронов, газ валентных электронов принадлежит «основной» зоне энергетических уровней и ряду расположенных выше «гармонических» зон. В атомистической модели валентные электроны находятся, кроме случаев перекрытия зон, в валентной зоне, которая возникает из валентного уровня изолированных атомов, в то время как расположенные выше уровни возникают из расположенных выше энергетических уровней изолированных атомов. Таким образом,

оказывается, что эти две картины строения зон приводят разными путями к одному и тому же результату. Мы не ставим своей целью показать особенности этих концепций в деталях и лишь отметим, что с позиций любой из этих двух теорий можно выяснить важнейшие черты зонной структуры электронных состояний в кристаллических твердых телах.

15-3. ИТОГИ

В кристаллическом твердом теле движение орбитальных и валентных электронов составляющих его атомов характеризуется зонами энергетических уровней. Вообще говоря, для каждого энергетического уровня изолированного атома имеется соответствующая зона энергетических уровней в твердом теле, построенном из большого числа этих атомов. В каждой зоне имеется столько близко расположенных энергетических уровней, сколько имеется атомов в образце, и каждый энергетический уровень может быть занят двумя электронами. В целом, если атомный энергетический уровень полностью занят электронами, соответствующая зона энергетических уровней в кристалле также заполнена. (Отметим не только исключение для зоны $2p$ кислорода в закиси меди, но и основательную причину для этого исключения.) Энергетические зоны очень узки в случаях, соответствующих энергиям внутренних орбит атомов. Зоны, соответствующие уровням валентных электронов и незанятым уровням с более высокой энергией, чем энергия валентных электронов, шире (несколько электронвольт от верхнего до нижнего края), и в пределах этих зон отдельные энергетические уровни расположены на расстоянии порядка 10^{-22} *эв* друг от друга в кристалле размером порядка 1 *см*. В некоторых веществах зоны валентных электронов перекрывают одну или большее число соседних зон.

Электроны в этих зонах движутся, подчиняясь законам квантовой механики, вследствие чего электроны на энергетических уровнях вблизи вершины зоны ведут себя так, как если бы их массы были отрицательными. Удаление одного из этих электронов ведет к образованию свободного места, которое ведет себя, как частица обладающая положительным зарядом и положительной массой. Такое свободное место является положительной дыркой.

ЗАДАЧИ

1. Основываясь на идеях квантовой механики, изложенных в тексте, покажите, что радиус самой глубокой внутренней разрешенной орбиты атома водорода равен $5{,}3 \cdot 10^{-9}$ *см*.

2. Докажите, что скорость электрона на этой орбите равна $2{,}2 \cdot 10^{8}$ *см/сек*.

3. Какова полная энергия водородного атома в электронвольтах, если его электрон находится на этой орбите?

4. Электрон заключен в потенциальный ящик длиной 1 *см*. Вычислите наименьшую квантованную скорость электрона, если он вынужден двигаться в этом пространстве размером 1 *см*.

5. Докажите, что если в потенциальном ящике заключено $5 \cdot 10^{7}$ электронов (это одномерный эквивалент кристалла в 1 $см^3$, содержащий $1{,}25 \cdot 10^{23}$ свободных валентных электронов) и эти электроны занимают уровни с наименьшими возможными для них импульсами по два электрона на уровень, то скорость и кинетическая энергия электрона у вершины этого распределения равны $9{,}15 \times 10^{7}$ *см/сек* и 2,3 *эв* соответственно.

ЛИТЕРАТУРА

1. Krutter H. M., Energy Bands in Copper, Physical Review, 1935, vol. 48, p. 664.

2. Herrings C., Transport Properties of a Many-Valley Semiconductors, Bell System Techn. Journal, 1955, vol. 34, 237.

3. Жданов Г. С., Физика твердого тела, Изд. Московского гос. университета, 1961.

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

Kittel C., Introduction to Solid-State Physics, 2nd ed., Wiley, New York, 1953, chapters 10, 11.

Seitz F., Physics of Metals, McGraw-Hill, New York, 1946, chapters 16, 17.

Seitz F., The Modern Theory of Solids, McGraw-Hill, New York, 1940, chapters 8, 13.

Shockley W., Electrons and Holes in Semiconductors, Van Nostrand, New York, 1950, chapters 5, 14.

Slater J. C., Quantum Theory of Matter, McGraw-Hill, New York, 1951, chapter 10.

ГЛАВА ШЕСТНАДЦАТАЯ

МЕХАНИЗМЫ ПРОВОДИМОСТИ ТВЕРДЫХ ТЕЛ

В предыдущей главе мы развили идею электронных энергетических зон в кристаллических твердых телах и показали, как на основе этой концепции может быть объяснено прохождение электричества через металлы. Мы также немного глубже проникли в природу положи-

тельно заряженных дырок. В результате теперь мы подготовлены к обсуждению в подробностях поведения электронов и дырок в этих энергетических зонах и можем, таким образом, аргументировать наше представление о явлениях электропроводности, имеющих место в твердых телах.

В иллюстрации к этой главе приняты несколько упрощенные изображения электронных энергетических зон. Из-за того, что мы редко сталкиваемся с поведением электронов во внутренних орбитах атомов твердого

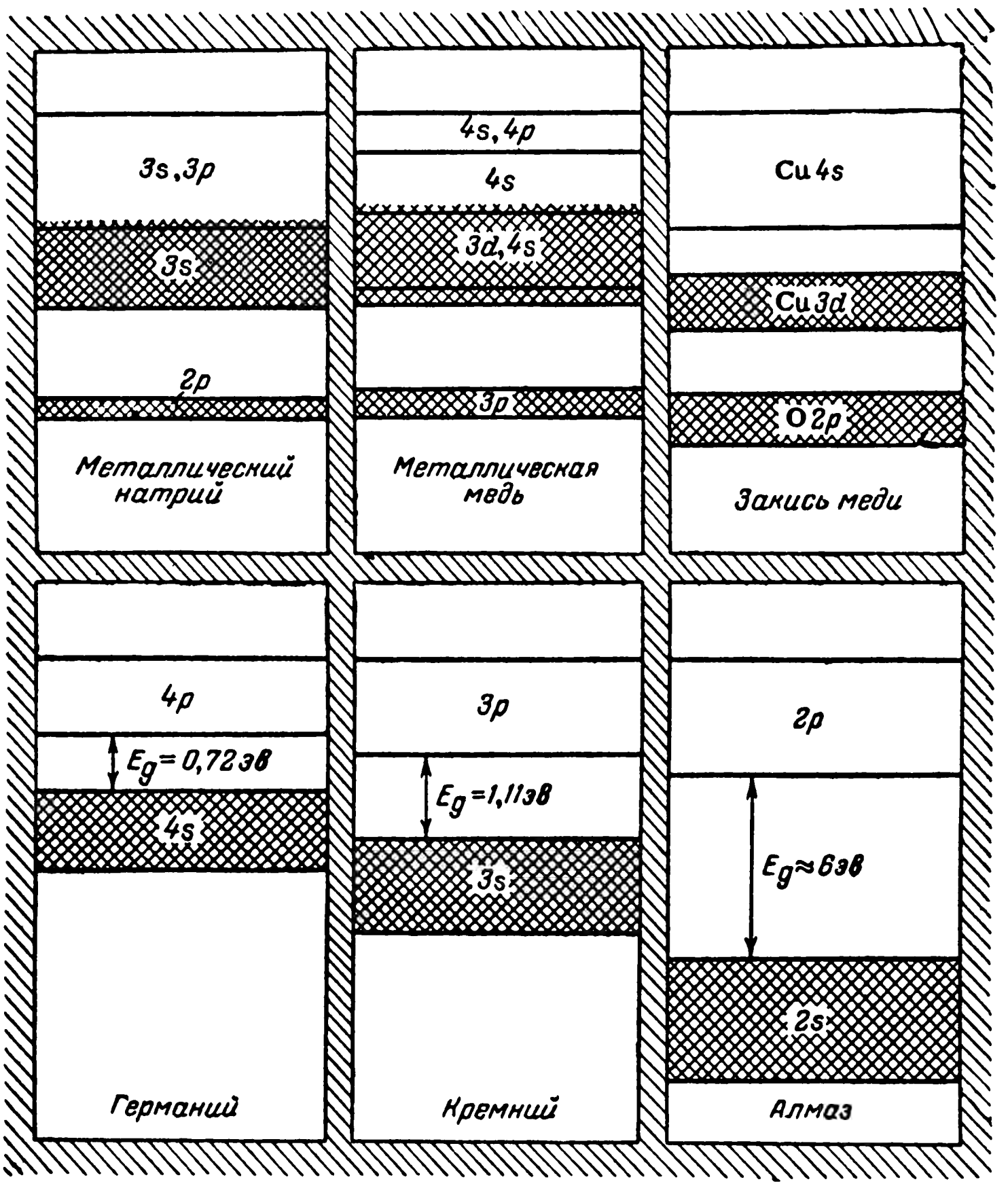


Рис. 16-1. У различных веществ схемы электронных энергетических зон различны, чем и объясняется разное электрическое поведение их

тела, мы не будем изображать на наших рисунках нижние энергетические зоны, а будем изображать лишь зону валентных электронов и ближайшие к ней или перекрывающиеся с нею зоны. Только поведение частиц в этих зонах интересно для наших целей и определяет свойства полупроводниковых материалов. Далее, мы не будем изображать отдельные энергетические уровни внутри зон и применим перекрестную штриховку, чтобы показать, какие зоны заполнены и в какой степени. Кроме особых случаев, когда это необходимо для аргументации, не будут изображаться также потенциальные горбы между атомами. На рис. 16-1 показаны упрощенные таким образом схемы энергетических зон для металлов — натрия и калия, для полупроводников — закиси меди, германия и кремния и для диэлектрика — алмаза.

Сразу становится очевидным различие между ними — это различный внешний вид схем энергетических зон для металлов, полупроводников и изоляторов. У металлов зоны частично заполнены; зоны полупроводников и изоляторов, с другой стороны, либо заполнены, либо совсем свободны. Затем, у изолятора энергетическая щель между верхней заполненной зоной и следующей расположенной выше свободной зоны шире, чем у полупроводников. Эти примеры очень хорошо иллюстрируют основные различия между металлами, полупроводниками и изоляторами, связанные с размещением и заполненностью энергетических зон. Посмотрим теперь, как эти различия приводят к возникновению наблюдаемых чрезвычайных различий в электропроводности этих трех классов твердых тел.

16-1. СОБСТВЕННАЯ ПРОВОДИМОСТЬ ПОЛУПРОВОДНИКОВ

Мы уже видели, как возникает металлическая проводимость в частично заполненной зоне с незанятыми энергетическими уровнями, на которые электроны могут переходить с помощью приложенного электрического поля, нарушающего симметрию движения электронов и создающего результирующий ток. Мы видели, что это нарушение равновесия достигает некоторой установившейся величины вследствие действия восстанавливающих факторов в виде связанных с потерей энергии столкновений и рассеяния, влияющих на электроны, перешедшие на уровни выше обычно занятых. Результи-

рующий ток в металле, таким образом, создается только небольшой частью электронов из тех, что заполняют зону, только теми электронами, которые находятся у вершины занятой электронами части зоны. Эта картина до некоторой степени находится в противоречии с классической картиной проводимости металла, согласно которой каждый валентный электрон ведет себя независимо от других и вносит свой вклад в результирующий ток. Мы также видели, что электроны *заполненной зоны* не могут участвовать в создании тока, так как их движения строго предопределены законами квантовой механики и не имеется незанятых уровней, в которые могли бы перейти электроны под действием приложенного электрического поля, нарушив распределение этих их движений. В такой заполненной зоне результирующее движение электронов симметрично, и ничем нельзя нарушить эту симметрию. Также, конечно, и совсем свободная зона не может способствовать проводимости, потому что в ней нет электронов, которые могли бы создать ток. Мы приходим поэтому к выводу, что *электропроводность может иметь место только за счет тех зон, которые частично заполнены или же частично свободны.*

Этого объяснения достаточно для понимания того, почему металлы проводят, а диэлектрики изолируют, но оно не объясняет, почему полупроводники — закись меди, кремний и германий, которые на рис. 16-1 имеют только заполненные или же совсем свободные зоны,— обладают известной нам из опытов проводимостью. На этих рисунках мы опять слишком упростили факты, стремясь выделить основные идеи концепции проводимости. Недостающая деталь заключается в том, что при всех температурах выше абсолютного нуля на нижних уровнях номинально свободных зон имеется некоторое количество электронов. Эти электроны переходят с верхних уровней номинально заполненных зон на нижние уровни свободных зон в результате термического возбуждения. При всех температурах атомы твердого тела колеблются около своих нормальных положений в кристаллической решетке, и чем выше температура, тем сильнее эти колебания. Время от времени энергия этих колебаний сообщается какому-либо электрону, в результате чего его полная энергия оказывается доста-

точной для скачка из заполненной нижней зоны в пустую верхнюю зону. Этот процесс приводит к тому, что в верхней зоне оказывается все время некоторое количество электронов: большее — при высоких температурах и меньшее — при низких. Этот процесс переброса, называемый термическим возбуждением, влияет в небольшой степени на относительную заполненность верхней зоны и освобожденность нижней зоны (рис. 16-2), создавая таким образом некоторую ограниченную прово-

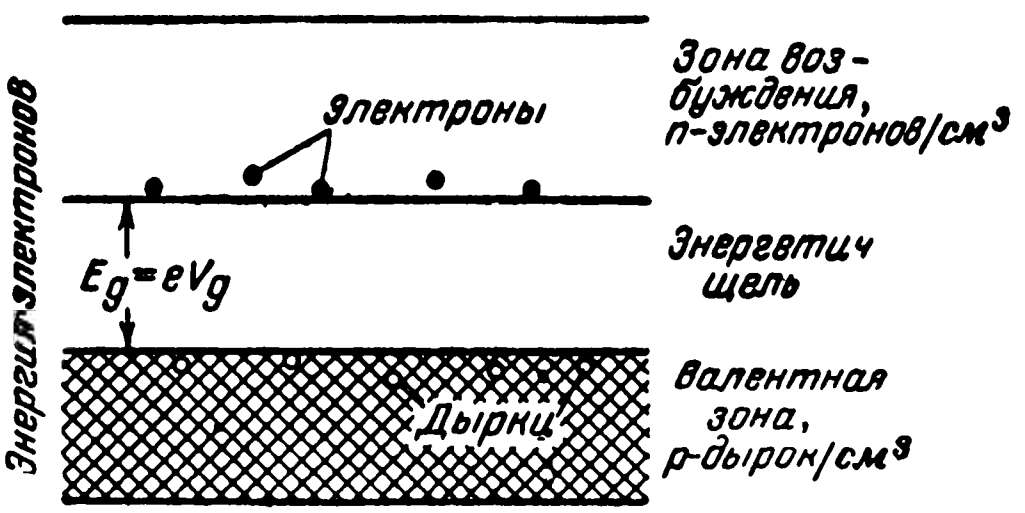


Рис. 16-2. Тепловое возбуждение поддерживает концентрации свободных электронов и дырок в чистых полупроводниках

димость за счет обеих зон. При переходе электрона в верхнюю зону он оставляет положительную дырку в нижней зоне. Электрон и дырка теперь становятся частицами-носителями заряда, могущими ускоряться электрическим полем и создавать ток. В верхней зоне появляется при этом столько же электронов, сколько дырок образуется в нижней зоне, и проходящий через образец ток оказывается созданным частично переносом электронов, частично — переносом дырок.

Только что описанный механизм есть механизм собственной проводимости полупроводников. Именно таково свойство полупроводникового образца, который либо достаточно чист, чтобы влияние оставшихся в нем примесей было пренебрежимо малым, либо поддерживается при достаточно высокой температуре, вследствие чего влияние примесей перекрывается собственной проводимостью. Число n электронов в кубическом сантиметре, переходящих вследствие нагревания в верхнюю зону, зависит от температуры T и ширины энергетической

щели E_g между двумя зонами [1] дозволенных состояний:

$$n \sim e^{-E_g/2kT} \sim e^{U_g/2kT}, \qquad (16\text{-}1)$$

где e — заряд электрона ($1{,}6 \cdot 10^{-19}$ *к*) и k — постоянная Больцмана ($1{,}4 \cdot 10^{-23}$ *дж/град*).

Возбужденные так электроны живут в верхней зоне недолго (у типичных полупроводников — несколько микросекунд), а затем рано или поздно рекомбинируют с положительными дырками в нижней зоне. Число n в действительности является результатом равновесия между процессами теплового возбуждения и рекомбинации, причем оба процесса все время идут с одинаковой быстротой.

Так как проводимость материала зависит непосредственно от числа участвующих в этом процессе частиц, то очевидно, что уравнение (16-1) дает выраженную через основные физические параметры экспоненциальную зависимость от температуры удельного сопротивления собственного полупроводника, которая раньше изображалась нами в виде эмпирического уравнения (2-1).

Таким образом,

$$\rho = \frac{1}{\sigma} \sim \frac{1}{n} \sim e^{eU_g/2kT} \sim e^{B/T}. \qquad (16\text{-}2)$$

Значения B, которые мы использовали в гл. 2 и 3 для характеристики этой температурной зависимости, теперь, как видно, определяются шириной энергетической щели U_g:

$$B = \frac{eU_g}{2k}. \qquad (16\text{-}3)$$

Используя это выражение, можно определить расстояние между зонами U_g для любого полупроводника в результате соответствующего расчета с измеренной величиной B в температурной области, где полупроводник ведет себя, как собственный.

Если сравнить энергетическую схему алмаза со схемами кремния и германия (рис. 16-1), то легко объяснить, почему алмаз—изолятор. Ширина энергетической

---

[1] В уравнении (16-1) E_g — энергия (в джоулях), необходимая для перехода электрона из нижней зоны в верхнюю Ширина этой энергетической щели может быть также выражена эквивалентным напряжением U_g *в*.

щели E_g у алмаза такая большая, что вероятность термического возбуждения электронов с переходом в верхнюю зону на много порядков меньше, чем у кремния и германия, в результате чего n для алмаза согласно уравнению (16-1) равно лишь нескольким электронам на кубический километр. В общем изолятор просто является веществом, у которого нет частично заполненных энергетических зон, а ширина энергетической щели так велика, что она препятствует возбуждению электронов в заметном количестве, необходимом для создания проводимости такого порядка, какая встречается у полупроводников. Поднимая температуру изолятора, можно повысить его проводимость до величины, характерной для полупроводника, если кристалл раньше не расплавится, не сгорит или его кристаллическая структура не изменится. С другой стороны, если охладить полупроводник до достаточно низкой температуры, то его проводимость уменьшится до величины, характерной для изолятора.

16-2. ПРИМЕСНАЯ ПРОВОДИМОСТЬ В ПОЛУПРОВОДНИКАХ; ДОНОРЫ И АКЦЕПТОРЫ

Посмотрим теперь, как изменяется проводимость полупроводника, если в кристалле его растворены примеси. На рис. 16-3 показано двухмерное представление кристаллической решетки германия. Германий четырехвалентен, и в кристалле его каждое ядро германия вместе с крепко связанными с ним внутренними элек-

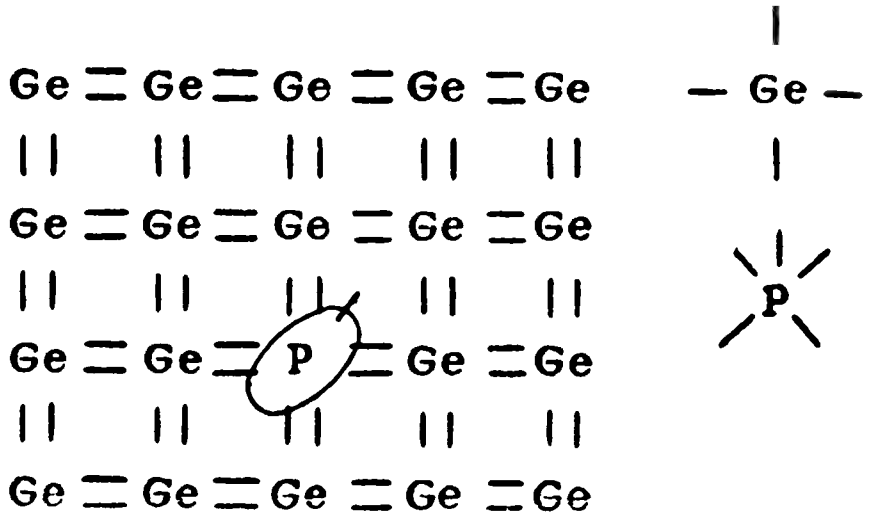


Рис. 16-3 Атом фосфора имеет в кристаллической решетке германия один избыточный валентный электрон. Короткие линии между атомами изображают валентные электронные связи между соседними атомами.

тронными оболочками окружено своими четырьмя валентными электронами. Этими электронами оно делится с четырьмя ближайшими соседями, получая от них в свою очередь взаймы по одному электрону от каждого из них. Поэтому каждый атом германия окружен восемью валентными электронами — четырьмя своими и четырьмя соседними. Взаимодействие этих обменивающихся валентных электронов создает силы, которые соединяют кристалл в единое целое. Эти электроны как

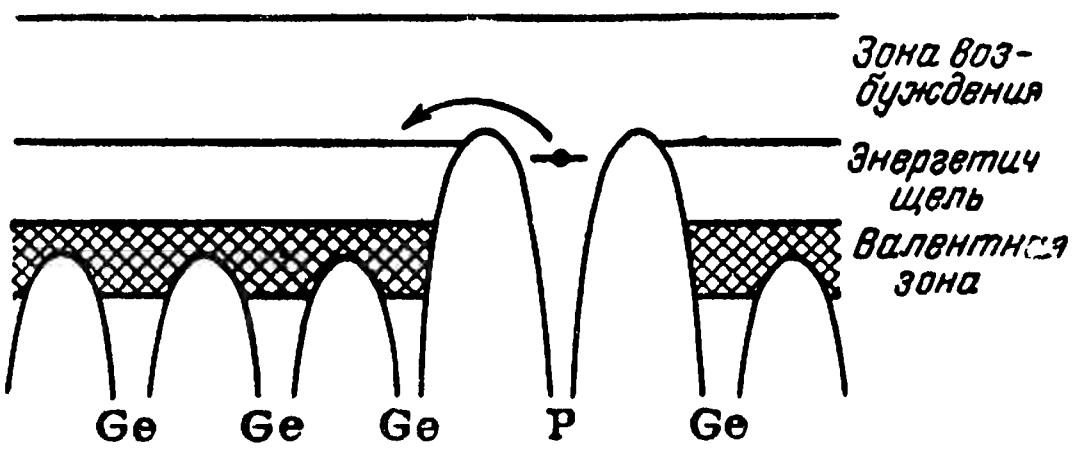


Рис 16-4 В атоме фосфора, занимающего место в кристаллической решетке германия, пятый валентный электрон занимает атомоподобный энергетический уровень в потенциальном колодце атома фосфора.

раз заполняют валентную зону атома германия, обозначенную 4*s* на рис. 16-1.

Если теперь, как показано на рис. 16-3, один атом какого-либо пятивалентного вещества, такого, как фосфор, заменяет в качестве примеси один атом в кристаллической решетке германия[1], четыре из пяти валентных электронов атома фосфора вступят во взаимодействие с четырьмя соседними атомами германия. Однако для пятого валентного электрона фосфора не находится места в плотном замкнутом объединении валентных электронов в кристалле. В результате этот «осиротевший» электрон обращается вокруг ядра фосфора на своей особой орбите, на которой он непрочно удерживается у атома. Все изложенное выше можно представить в виде картины энергетических уровней (рис. 16-4).

[1] Атомы фосфора слишком велики, чтобы поместиться в промежутках между атомами кристалла германия. Единственная возможность существования фосфора в качестве атомарно распределенной примеси в германии — замещать последний в его регулярной решетке.

Атом фосфора представляется в виде потенциальной ямы, которая удерживает пятый валентный электрон на локализованном энергетическом уровне как раз под нижним краем верхней зоны. Пока электрон остается на этом уровне, он «принадлежит» атому фосфора, но требуется совсем невысокая тепловая энергия, для того чтобы переместить его в верхнюю зону кристалла

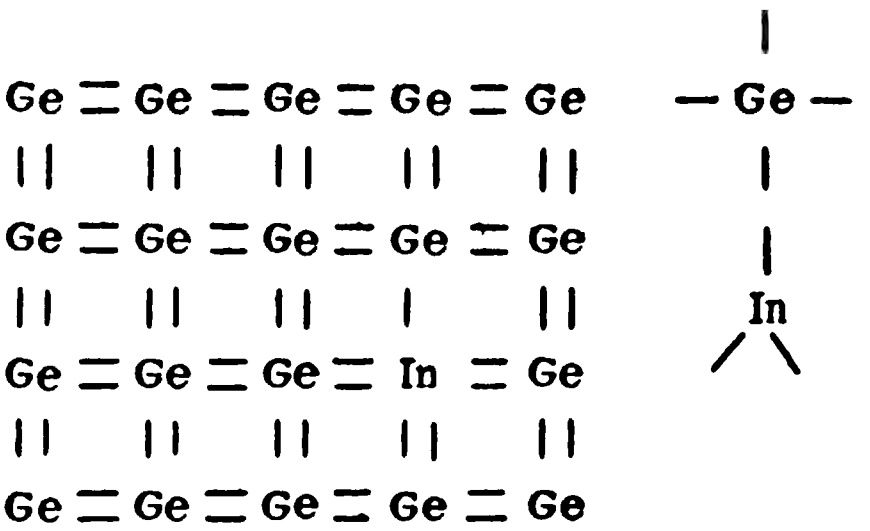


Рис. 16-5 Атом индия в кристаллической решетке германия создает дефицит одного валентного электрона, отсутствие которого может отразиться на соседних атомах германия.

германия, где он становится свободным электроном, т. е. электроном проводимости.

Проводимость полупроводника может быть, таким образом, сильно увеличена и может намного превышать собственную проводимость в результате появления этих дополнительных электронов проводимости, отданных подобными примесями. Примеси, обладающие бóльшим числом валентных электронов, чем требуется для связи в кристаллической решетке, называются дóнорами, так как они отдают дополнительные электроны в свободную зону. Присутствие донорных примесей в полупроводнике приводит к появлению проводимости типа *n* с концентрацией свободных электронов в зоне возбуждения, превышающей концентрацию положительных дырок в валентной зоне.

Рассмотрим последствия растворения примеси другого вида в полупроводниковом кристалле. На рис. 15-5 показан кристалл германия с атомом индия, замещающим один атом германия. Индий трехвалентен и его атом не обладает достаточным числом электронов, чтобы полностью заместить все валентные связи в кристал-

лической решетке. Однако под действием сравнительно небольшой энергии теплового возбуждения он может отнять электрон у одного из своих соседей. Ограбленный атом в свою очередь может отнять электрон у одного из его соседей и т. д. Недостаток в электроне перемещается, таким образом, от атома к атому в виде положительной дырки, а вызванное этим частичное

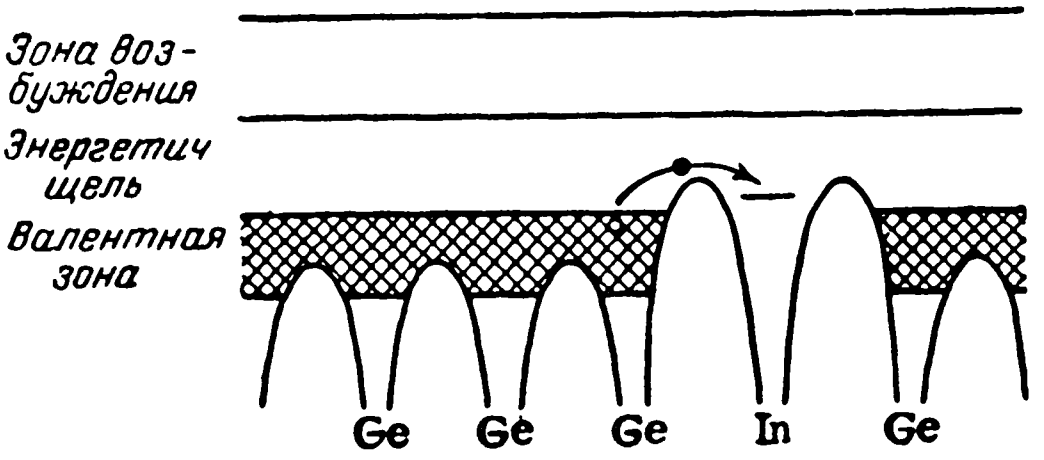


Рис 16-6 Потенциальная яма атома индия, занимающего место в кристаллической решетке германия, содержит свободный энергетический уровень, на который электрон может быть переброшен из нижней зоны тепловым возбуждением, оставив в валентной зоне дырку

опустошение валентной зоны позволяет ей участвовать в создании электропроводности.

Разъяснение этого процесса с помощью картины энергетических уровней дано на рис. 16-6. Атом индия представлен в виде потенциальной ямы со свободным локализованным энергетическим уровнем, расположенным немного выше верхнего края валентной зоны. При термическом возбуждении электрон из валентной зоны может перескочить на этот уровень, оставив в валентной зоне кристалла германия положительную дырку. Примеси, которые захватывают подобным образом электроны из валентной зоны, называются акцепторами; их присутствие в полупроводниковом образце приводит к появлению проводимости *p*-типа.

В полупроводниковом образце, содержащем доноры и (или) акцепторы, происходит непрерывное движение, связанное с опустошением и заполнением энергетических уровней и обменом электронов, причем все это происходит в результате теплового возбуждения кристаллической решетки. Электрон может сделать скачок из донорного уровня, побродить некоторое время по решет-

ке, находясь на верхнем энергетическом уровне, затем снова упасть на пустой донорный уровень где-нибудь в кристалле, или же спрыгнуть в валентную зону, уничтожив в ней проходящую вблизи положительную дырку. Электроны из валентной зоны перепрыгивают на акцепторные уровни и обратно и т. д. При любой температуре все эти различные процессы происходят таким образом, что устанавливается динамическое равновесие, которое мы обсудим в следующей главе.

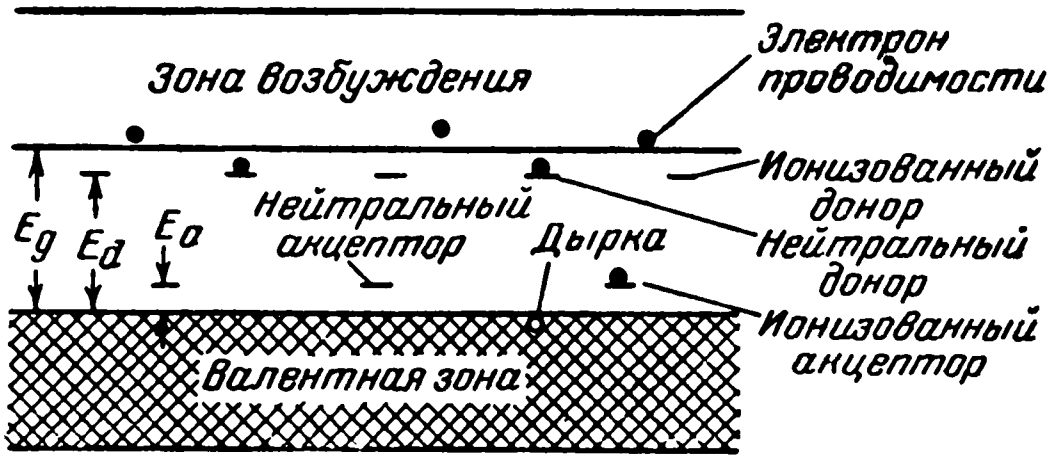


Рис. 16-7. Принятое изображение полупроводника, обладающего как донорами, так и акцепторами

В дальнейшем будет удобно чертить энергетические диаграммы, в которых донорные и акцепторные уровни представлены просто в виде коротких черточек на соответствующих уровнях этих диаграмм. Такая упрощенная диаграмма показана на рис. 16-7 для полупроводника, содержащего доноры и акцепторы, но с бóльшим количеством доноров, чем акцепторов. Такой образец полупроводника обнаружил бы в результате преобладания доноров свои свойства как вещество *n*-типа. На этом рисунке показаны различные обозначения величин, которые использованы при дальнейших рассуждениях.

Теперь легко понять, почему содержащий примеси полупроводник проявляет преимущественно примесную проводимость при низких температурах и собственную проводимость при высоких температурах. Рассмотрим, например, полупроводник, содержащий примеси только донорного типа. По мере повышения температуры от абсолютного нуля первые свободные электроны, которые должны появиться в зоне возбуждения, придут из донорных уровней, потому что энергия, требуемая для их освобождения, мала. По мере повышения температуры возбуждается все больше электронов и проводи-

мость n-типа этого образца возрастает с ростом температуры. Однако, начиная с некоторой температуры, от доноров нельзя уже больше получить электронов из-за истощения донорных уровней и ограниченности числа доноров. Одновременно с этим при повышении температуры электроны начинают поступать из валентной зоны и все в большем количестве. Так как валентная зона является почти неистощимым источником электронов, чис-

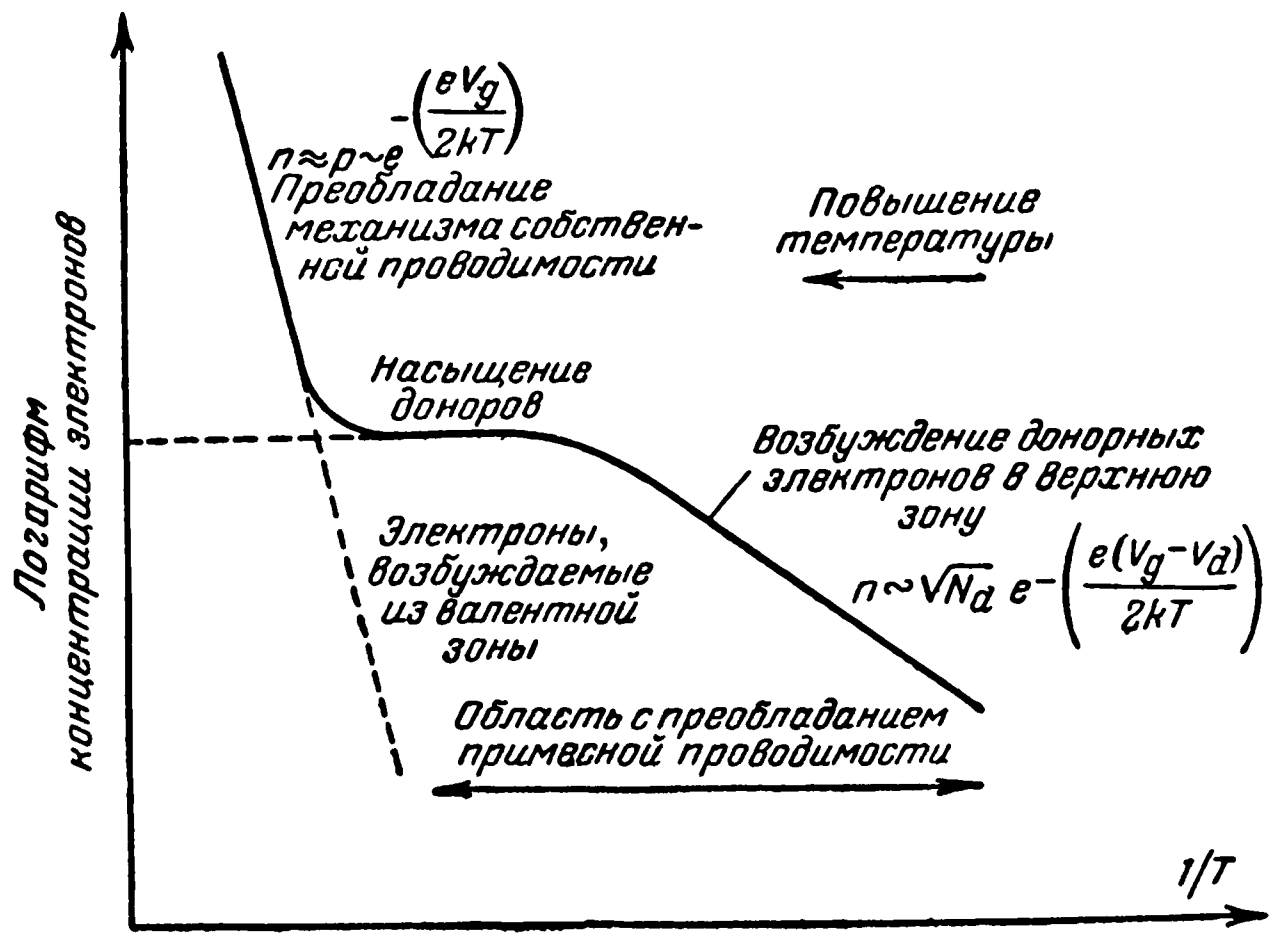


Рис. 16-8. Концентрация электронов в зоне возбуждения полупроводника n-типа зависит от температуры Сравните это с кривыми электропроводности на рис. 2-5.

ло возбужденных электронов (и число положительных дырок, которые они оставляют после себя) может превысить число донорных электронов. Если это произойдет, образец начнет проявлять собственную проводимость, которая перекрывает примесную проводимость и становится все большей по мере того, как температура поднимается дальше. На рис. 16-8 представлена кривая числа электронов в верхней зоне на кубический сантиметр примесного полупроводника n-типа в виде функции температуры. Чтобы легче было сравнить эту кривую с кривой зависимости проводимости от температуры, изображенной на рис. 2-5 и 2-6, здесь представлен логарифм n в зависимости от $1/T$. Две пунктирные кривые указывают отдельно вклад от возбуждения доноров и от возбуждения электронов основ-

ной решетки (собственного возбуждения). Сплошная кривая дает полную концентрацию электронов, складывающуюся из концентраций обоих типов. Все рассуждения, приведенные выше для полупроводника *n*-типа, конечно, в такой же степени применимы и к полупроводнику *p*-типа.

Сходство между кривой на рис. 16-8 и кривыми графиков зависимости логарифма σ от $1/T$ для германия и закиси меди в гл. 2 очевидно. Ясно, что изменение проводимости в зависимости от температуры в этих полупроводниках должно определяться изменением концентрации свободных носителей заряда от температуры. Однако полное уравнение для проводимости σ имеет вид:

$$\sigma = n\mu_n e + p\mu_p e, \qquad (16\text{-}4)$$

где n — число электронов в 1 $см^3$ в зоне возбуждения;
p — число дырок в 1 $см^3$ в валентной зоне;
μ_n и μ_p — подвижности электронов и дырок, соответственно;
e — заряд электрона или дырки.

Чтобы закончить описание механизма проводимости, необходимо рассмотреть подвижности электронов и дырок и зависимость их от температуры.

16-3. ПОДВИЖНОСТЬ И УДЕЛЬНАЯ ЭЛЕКТРОПРОВОДНОСТЬ

Подвижность заряженной частицы есть средняя установившаяся направленная скорость, которую она приобретает в направлении приложенного электрического поля напряженностью 1 *в/см* (скорость дрейфа). Действительный путь частицы очень изломан, так как он постоянно прерывается столкновениями, которые частица испытывает с атомами решетки, с другими частицами, с примесями и включениями, с поверхностями образца и т. п. Таким образом, подвижность — величина статистическая; для общего числа частиц ее можно рассматривать как константу материала, зависящую от температуры и содержания примеси. Подвижности выражаются в сантиметрах в секунду на вольт на сантиметр.

Электроны проводимости полупроводника находятся на нижних энергетических уровнях зоны возбуждения; положительные дырки проводимости находятся на верх-

них уровнях валентной зоны. Обычно этих частиц так мало по сравнению с числом доступных для них энергетических уровней, что не в пример случаям полузаполненных или совсем заполненных энергетических зон здесь почти нет ограничений поведения частиц, вызванных наличием частиц на соседних уровнях. Поэтому участвующие в проводимости частицы, находящиеся в почти идеальном тепловом равновесии с кристаллической решеткой, имеют практически максвелловское распределение по скоростям и участвуют вместе с решеткой кристалла в равнораспределении тепловой энергии. Поэтому средняя тепловая кинетическая энергия частицы, участвующей в проводимости, равна $^3/_2\ kT$ *.

Траектория единичного электрона или дырки, если рассматривать ее с корпускулярной, а не волновой точки зрения, представляет собой последовательность прямых отрезков между точками столкновений. Длина любого из этих прямых отрезков, определенная как средняя из многих таких отрезков, является средним свободным пробегом. Средний промежуток времени между последовательными столкновениями частицы является средним временем свободного пролета. Можно очень просто получить соотношение между подвижностью частицы и средним свободным пробегом. Все время пока частица свободна, она может ускоряться электрическим полем. Это ускорение увеличивает скорость, которой частица уже обладает, и прирост скорости направлен в сторону поля. При следующем столкновении частица отдает решетке избыток импульса, который она получила со времени последнего столкновения. При этом необязательно при каждом столкновении передается весь приобретенный импульс. Частица может испытать несколько последовательных упругих столкновений, а затем передать накопленный излишек импульса решетке за один удар. Однако импульс частицы обычно не выходит далеко за пределы средней величины, соответствующей тепловому равновесию в обычных полях, так как всегда находится какой-нибудь ограничительный процесс, возвращающий его к величине, отвечающей равновесному или почти равновесному состоянию. Дви-

---

* Этот результат приводится без доказательства, которое относится к области кинетической теории.

жение частицы в этом поле происходит маленькими скачками, постоянно прерываемыми столкновениями, подобно движению монеты, падающей сквозь лабиринт из проволочек. Постепенное перемещение частицы создает ток.

На основе этой чисто классической картины легко подсчитать, что подвижность равна:

$$\mu = \frac{el}{2\sqrt{3mkT}} \frac{см^2}{эл.\text{-}стат.\ ед.\ потенц.\cdot сек}, \qquad (16\text{-}5)$$

где все физические величины выражены в электростатических единицах и единицах системы СГС.

Более точный вывод, учитывающий, что не все свободные пробеги имеют одинаковую длину, в котором осуществляется усреднение по всем длинам пробегов, вводит дополнительный коэффициент

$$\frac{4\sqrt{2}}{\sqrt{3\pi}},$$

что дает окончательно:

$$\mu = \frac{4el}{3\sqrt{2\pi mkT}} = \frac{el}{225\sqrt{2\pi mkT}} \frac{см^2}{в\cdot сек}. \qquad (16\text{-}6)$$

Примечательно, что это выражение не зависит от электрического поля, что свидетельствует о линейности тока по отношению к напряжению. Очевидно также, что подвижность пропорциональна среднему свободному пробегу и определяется частично природой механизма, ограничивающего средний свободный пробег. Очевидно также, что подвижность обратно пропорциональна квадратному корню из температуры, т. е.

$$\mu \sim \frac{l}{\sqrt{T}}. \qquad (16\text{-}7)$$

Теперь посмотрим, как наши теоретические предположения согласуются с действительностью. На рис. 16-9 представлены кривые зависимости подвижности[1] от температуры для положительных дырок в кремниевых

[1] Способы измерения подвижностей изложены в следующих главах.

образцах *p*-типа, активированных различными количествами бора [Л. 1]. Отметим, во-первых, что увеличение концентрации примеси приводит к понижению подвижности. Это явление согласуется с допущением, что примесные атомы, растворенные в кристаллической решетке, сокращают средние свободные пробеги подвижных частиц.

Однако труднее понять, почему кривые подвижностей, показанные на рис. 16-9, не следуют зависимости

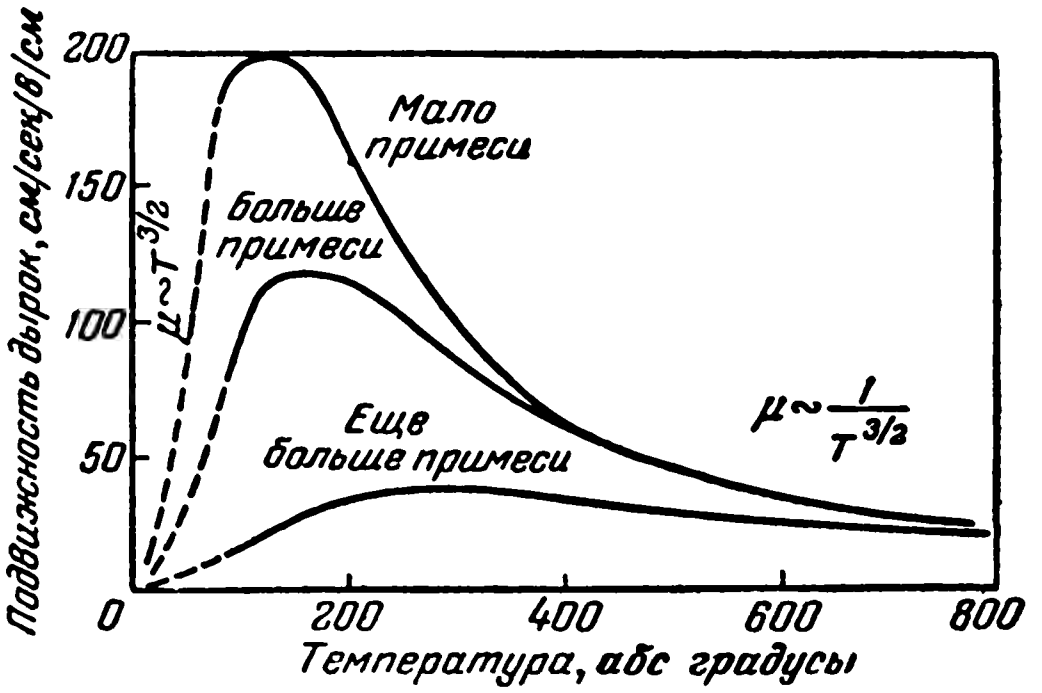


Рис. 16-9 Подвижность положительных дырок в кремнии *p*-типа зависит от температуры и концентрации примеси.

обратной пропорциональности квадратному корню из температуры согласно уравнению (16-7). При низких температурах, для которых экспериментальные кривые экстраполированы пунктирными линиями, подвижность, по-видимому, пропорциональна полуторной степени температуры, хотя из-за грубости в экстраполяции точная величина этого показателя не определена. При высоких температурах, по-видимому, все кривые графика сходятся к общей асимптотической кривой подвижности, характерной для всех образцов чистого кремния и обратно пропорциональной температуре в степени три вторых.

Таким образом,

$$\left.\begin{aligned} &\text{при низкой температуре } \mu \sim T^{3/2}; \\ &\text{при высокой температуре } \mu \sim \frac{1}{T^{3/2}}. \end{aligned}\right\} \qquad (16\text{-}8)$$

Это расхождение теоретических данных, предсказываемых уравнением (16-7), с данными эксперимента, выражаемыми уравнениями (16-8), можно устранить, допустив, как это ни покажется нелогичным, что сам средний свободный пробег зависит от температуры следующим образом:

$$\left.\begin{array}{l}\text{при низких температурах } l \sim T^2; \\ \text{при высоких температурах } l \sim \frac{1}{T}.\end{array}\right\} \qquad (16\text{-}9)$$

Почему же средний свободный пробег столь по-разному ведет себя при разных температурах? Для ответа на этот вопрос мы должны исследовать механизмы, которые ограничивают средний свободный пробег. Делая это, мы используем как корпускулярную физику, так и волновую физику, применяя ту из них, которая дает лучшие доказательства.

При низких температурах механизмом, ограничивающим средний свободный пробег, являются столкновения между свободными частицами и атомами примеси, границами кристалла, свободными ячейками в решетке или другими искажениями идеальной решетки кристалла. Удивительным фактом является то, что столкновения с атомами кристаллической решетки (самыми многочисленными атомами в нашем случае) несущественны при очень низких температурах. Чтобы выяснить, почему это происходит, мы должны рассмотреть не сами электроны, а перемещающиеся в кристалле электронные волны. При очень низких температурах расположение атомов в кристалле очень близко к идеальному порядку, представляя собой среду для распространения электронных волн с почти идеальной периодичностью. Волна *через идеальную периодическую структуру* проходит без рассеяния, отражения или потерь, если структура нерассеивающая и если длина волны соответствует разрешенной частоте в зоне пропускания структуры. В такой среде на распространение волны не влияет зернистость самой структуры. Таким образом, *в идеальном кристалле вблизи абсолютного нуля температуры* электронные волны совсем не «замечают» атомов решетки и средний свободный пробег в нем должен ограничи-

ваться только размерами самого образца[1]. Если в таком кристалле имеются посторонние атомы, свободные ячейки кристаллической решетки или же другие нарушения идеальной структуры, они становятся рассеивающими центрами для электронных волн. Кристаллическая решетка ведет себя в этом отношении аналогично передающей линии, а передающая линия, имеющая местные нарушения параметров распространения, отражает бóльшую или меньшую часть входной волны в зависимости от количества и величины этих нарушений. Когда электронная волна отражается или рассеивается таким рассеивающим центром, электрон рассеивается вместе с нею и мы говорим, что электрон столкнулся с этим центром. Средние свободные пробеги свободных электронов в кристалле ограничиваются, таким образом, этими столкновениями. У образцов с высокой концентрацией примесей рассеяние происходит часто, средние свободные пробеги малы, подвижности частиц невелики. Эти качества проявляются сильнее при увеличении количества примесей в образце (рис. 16-9) в области температур, где все свойства зависят от примесей. (Положительная дырка характеризуется волной Де-Бройля точно так же, как и электрон, и все приведенные выше рассуждения также подходят для положительных дырок и образцов *p*-типа.)

Теперь посмотрим, почему в образце, где примесное рассеяние является доминирующим механизмом ограничения среднего свободного пробега, последний должен быть пропорциональным квадрату температуры, как в уравнении (16-9)?

Чтобы увидеть, как получается эта зависимость, мы должны вспомнить, что при низких температурах свободные частицы собираются на нижних энергетических уровнях в своих соответствующих зонах и движутся с малыми тепловыми скоростями. Медленная частица, проходящая близко к отклоняющему центру, пребывает вблизи него достаточно долго, чтобы заметно отклониться в результате взаимодействия с центром. С другой стороны, при высоких температурах тепловая скорость частицы выше и она пролетает мимо центра так быстро,

---

[1] В настоящий момент мы не рассматриваем процессы, связанные с нулевыми колебаниями решетки.

что не отклоняется значительно. Столкновение, которое при низких температурах приводит к рассеянию, при высоких температурах практически не будет заметно. Поэтому эффективное сечение столкновения с рассеивающим центром уменьшается с увеличением температуры. Хотя в этом утверждении нет подтверждения зависимости $l \sim T^2$ вместо, скажем, зависимости вида $l \sim T$ или $l \sim T^{3/2}$, оно все же объясняет, почему средний сво-

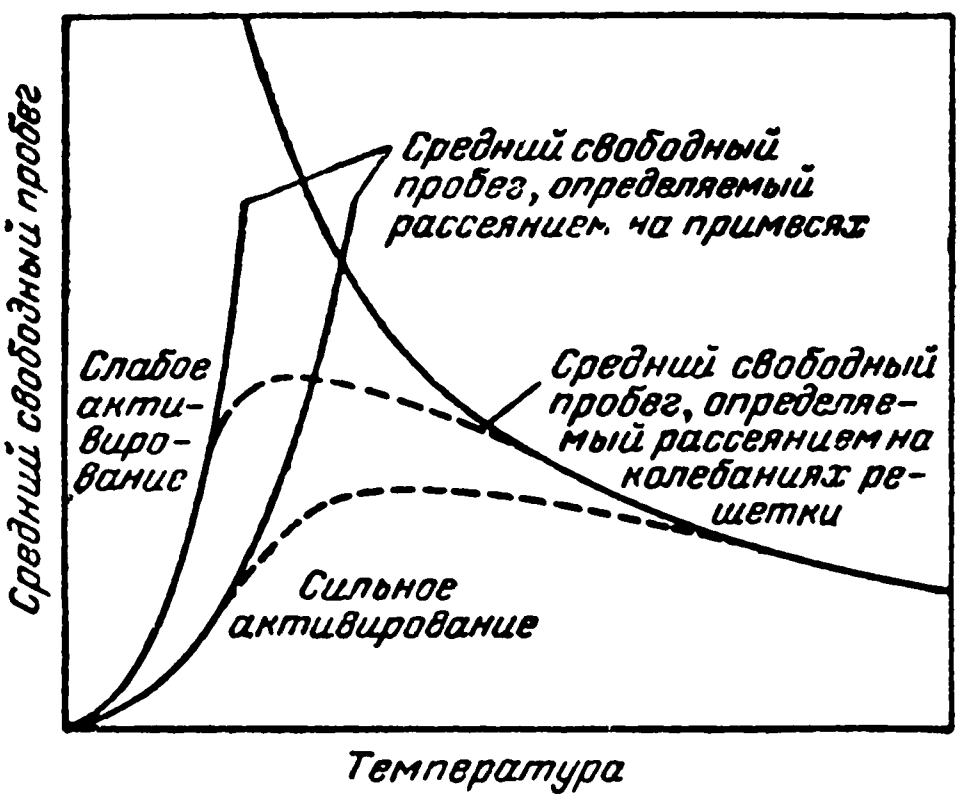


Рис. 16-10. Средний свободный пробег определяется концентрацией примеси и тепловыми колебаниями кристаллической решетки.

бодный пробег увеличивается с ростом температуры в области низких температур [Л. 2].

Это поведение можно иллюстрировать графически, например, как на рис. 16-10, где средний свободный пробег, определяемый рассеянием на примесях, представлен в виде кривых зависимости от температуры для двух образцов полупроводника с различными концентрациями активирующих примесей. Две кривые, отмеченные как «Средний свободный пробег, определяемый рассеянием на примесях», демонстрируют влияние этого уменьшения сечения соударения в зависимости от возрастающей температуры. Кривая на рис. 16-10, обозначенная как «Средний свободный пробег, определяемый рассеянием на основной решетке», иллюстрирует второй основной механизм рассеяния, сравнительно несущественный при очень низких температурах, но доми-

нирующий при температурах достаточно высоких, чтобы рассеяние на примесях стало несущественным. Этот механизм является рассеянием электронных волн тепловыми колебаниями кристаллической решетки. При всех температурах выше абсолютного нуля атомы кристалла вследствие тепловых колебаний смещены с их обычных мест в решетке, что с точки зрения электронных волн означает появление случайной составляющей в постоянной распространения решетки. Это обусловливает рассеяние, возрастающее с увеличением температуры и приводящее к постепенному уменьшению среднего свободного пробега [Л. 3]. Средний свободный пробег, определяемый этим механизмом, на рис. 16-10 представлен гиперболой «Рассеяние решеткой». Эта кривая характеризует само кристаллическое вещество и не зависит от содержания примеси в области примесных концентраций, представляющих интерес для технологии полупроводников.

Так как в любом реальном образце имеют место одновременно как примесное рассеяние, так и рассеяние тепловыми колебаниями решетки, то общее рассеяние определяется суммой этих двух рассеяний и результирующие средние свободные пробеги оказываются короче, чем в том случае, если бы действовал какой-нибудь один из этих механизмов. Результирующие средние свободные пробеги для двух гипотетических образцов представлены на рис. 16-10 пунктирными кривыми, которые внешне похожи на экспериментальные кривые зависимости подвижности от температуры, изображенные на рис. 16-9. Каждая из кривых на рис. 16-9 может быть преобразована в соответствующую кривую на рис. 16-10 делением ординат на $\sqrt{T}$, как это следует из уравнения (16-7), и соответствующим подбором констант.

Таким образом, мы показали, как подвижность свободных частиц в полупроводнике может быть объяснена на основе элементарных физических процессов. Мы привели здесь эти несколько сложные рассуждения только для того, чтобы глубже показать детали механизма проводимости.

Теперь нам ясно, что пики кривых подвижности на рис. 16-9 возникают при тех температурах, при которых с возрастанием температуры рассеяние на примесях затухает и основным становится рассеяние на тепловых

колебаниях решетки, определяющее средний свободный пробег.

В кремнии и германии с обычными концентрациями примесей эти максимумы подвижности лежат в пределах 100° С в ту или другую сторону от комнатной температуры.

Интересно теперь эти данные о подвижности учесть в зависимости проводимости от температур.

Входящие в уравнение

$$\sigma = n\mu_n e + p\mu_p e \qquad (16\text{-}4)$$

величины можно вычертить отдельно в функции от температуры и выполнить графически требуемые уравнения умножения и сложения для получения окончательной

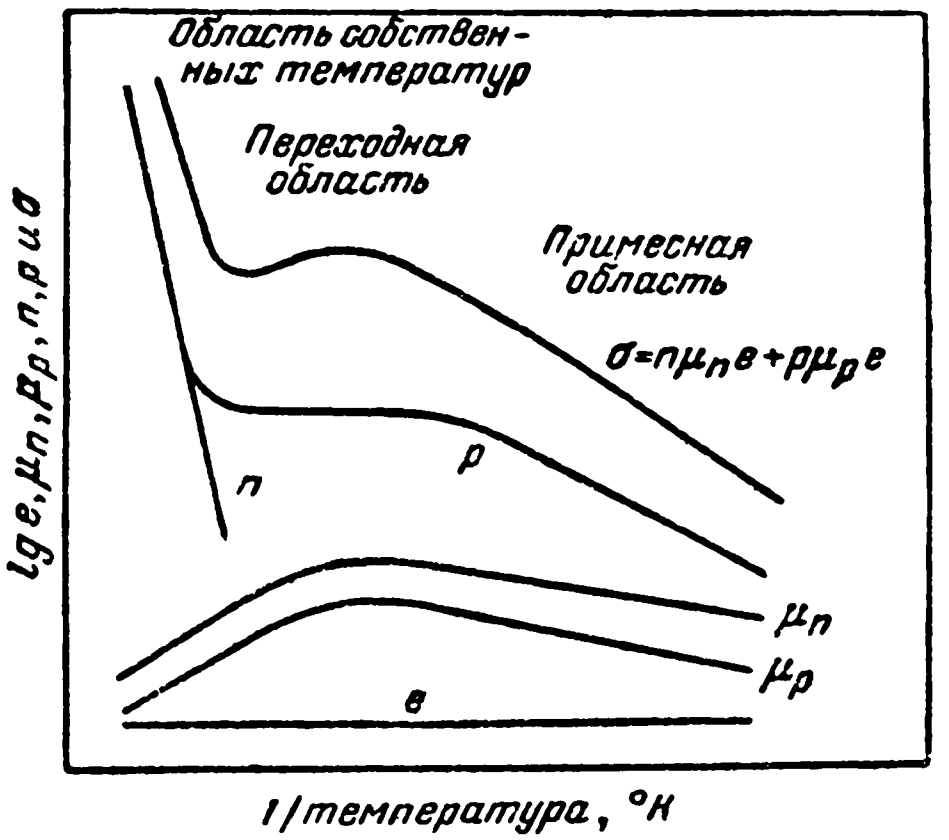


Рис. 16-11. Логарифм электропроводности $\log \sigma$ в зависимости от $1/T$ может быть рассчитан из зависимостей от $1/T$ величин n, p, μ_n, μ_p и e.

кривой зависимости σ от T. Это показано на рис. 16-11 для воображаемого полупроводника p-типа при низких температурах. На этом рисунке логарифмы n, μ_n, p, μ_p и e представлены в зависимости от $1/T$ для сравнения получающейся кривой зависимости проводимости от температуры с соответствующими кривыми на рис. 2-5 и 2-6. Сходство получается полное (за исключением

коэффициентов, определяющих относительные размеры и масштабы) вплоть до предсказания спада удельной проводимости германия в переходной температурной области. На кривых для закиси меди такого снижения не наблюдается, ввиду того, что при увеличении температуры собственная электропроводность устанавливается прежде, чем акцепторы истощаются, и переходный температурный интервал оказывается очень коротким. Каждая особенность этих кривых зависимости удельной электропроводности от температуры может быть объяснена с точки зрения этих основных физических свойств; в свою очередь изучение кривых зависимости удельной электропроводности от температуры для какого-либо частного полупроводника может дать нам сведения о ширине энергетической щели U_g, энергии возбуждения доноров и (или) акцепторов, о подвижности и температурных зависимостях этих величин.

Предыдущий анализ связи среднего свободного пробега с подвижностью был проведен в предположении, что прирост скорости, полученной носителем заряда от электрического поля, всегда невелик по сравнению с тепловой скоростью. В случае больших напряженностей поля это предположение не оправдывается. Частицы в среднем «горячее», чем решетка, в которой они движутся. Выравнивание распределения энергии между частицами и решеткой отстает от накопления частицами избытка энергии, и подвижность начинает зависеть от напряженности электрического поля. В этой области напряженностей электрического поля ток, вместо того чтобы возрастать линейно с напряжением, возрастает пропорционально квадратному корню из напряжения. В кремнии и германии возникновение такой аномальной зависимости удельной электропроводности происходит при напряженности электрического поля приблизительно около 1 000 *в/см* при комнатной температуре.

ЗАДАЧИ

1. Определите ширину запрещенных зон для кремния, германия и закиси меди, используя график зависимости удельного сопротивления от температуры на рис. 2-4.

2. Покажите, что удельное сопротивление чистого алмаза при комнатной температуре должно быть $2{,}7 \cdot 10^{46}$ *ом · см*, если для алмаза $U_g = 6$ *в* и $\mu_n = \mu_p = 1\,000$ *см²/сек · в*.

3 Выведите простое классическое уравнение подвижности (16-5). Считайте, что все участвующие в проводимости частицы имеют одинаковую тепловую кинетическую энергию $3kT/2$ и что при каждом столкновении каждая частица отдает решетке всю избыточную кинетическую энергию, которую она получила от электрического поля во время предыдущего свободного пробега.

4. Образец кремния p-типа, активированный акцепторной примесью до состояния, соответствующего средней кривой на рис. 16-9, имеет удельное сопротивление 0,03 *ом · см* при комнатной температуре. Какова в нем концентрация положительных дырок? В этой задаче можно не учитывать влияния электронов на проводимость.

5 Образец германия p-типа с удельным сопротивлением 48 *ом · см* и концентрацией положительных дырок $5 \cdot 10^{13}$ *см*$^{-3}$ облучается вспышкой света, что освобождает в нем еще по $2 \cdot 10^{13}$ электронов и дырок в каждом кубическом сантиметре. Покажите, что концентрация электронов в образце до освещения была $1{,}25 \cdot 10^{13}$ *см*$^{-3}$ и что после вспышки удельное сопротивление упало до 26,5 *ом · см*. Примите $\mu_p = 1\,700$; $\mu_n = 3\,600$ *см*^{2}/*сек · в*.

6. Так как подвижности электронов и дырок в германии различны, то самое высокое возможное удельное сопротивление не соответствует точному равенству электронной и дырочной концентраций Вычислите отдельно: а) собственное удельное сопротивление и б) максимальное возможное удельное сопротивление. В какую сторону (к преобладанию электронной или дырочной проводимости) отклоняется состояние от точного равенства в последнем случае чистого полупроводника?

ЛИТЕРАТУРА

1. Pearson G. L., Bardeen J., Electrical Properties of Pure Silicon and Silicon Alloys, Physical Review, 1949, vol 75, p. 865.

2 Conwell E. a. Weisskopf V F, Theory of Impurity Scattering in Semiconductors, Physical Review, 1950, vol. 77, p. 388.

3. Shockley W., Electrons and Holes in Semiconductors, Van Nostrand, 1950, chapter 11.

Русский перевод Шокли В., Теория электронных полупроводников, приложения к теории транзисторов, Изд. иностранной литературы, 1953.

4. Ryder E. J., Shockley W., Interpretation of Dependence of Resistivity of Germanium on Electric Field, Physical Review, 1949, vol. 75, p. 310.

5. Shockley W., Hot Electrons in Germanium and Ohm's Law, Bell System Technical Journal, 1951, vol. 30, p. 779.

6. Ryder E. J., Mobility of Holes and Electrons in High Fields, Physical Review, 1953, vol. 90, p. 766.

7. Аваkьянц Г. М., Феменологическая теория полупроводников, Изд. АН УзССР, Ташкент, 1960.

8. Иоффе А. Ф., Физика полупроводников, Изд. АН СССР, 1957.

ГЛАВА СЕМНАДЦАТАЯ

СТАТИСТИКА ЭЛЕКТРОНОВ И ДЫРОК

17-1. ВВЕДЕНИЕ

В этой главе мы исследуем с помощью статистики некоторые аспекты свойств электронов и дырок в твердых телах. Для понимания зависимости удельной электропроводности полупроводников от температуры важно ясно понимать, чем определяется зависимость числа частиц, определяющих электропроводность, от температуры. Мы уже видели, что в примесном полупроводнике удельная электропроводность растет по мере повышения температуры до момента, когда доноры истощаются (или акцепторы насыщаются). Однако при некоторой температуре начинает преобладать собственная электропроводность, маскирующая особенности примесной электропроводности и приводящая к дальнейшему росту удельной электропроводности с увеличением температуры.

Рассмотрим теперь математическое описание этих процессов, получаемое с помощью статистики Ферми—Дирака. Мы при этом не только найдем способ рассчитывать концентрации носителей зарядов как функции температуры, но разовьем также представление об уровне Ферми, которое широко использовано в дальнейшем при рассмотрении теории p-n перехода, теории выпрямления, а также теории фотоэлектрических и термоэлектрических явлений.

По одному из основных законов природы система, предоставленная самой себе, будет перестраиваться, стремясь к самому низкому энергетическому состоянию. Так, вода оказывается на дне сосуда, заряд течет по проводнику к более низкому потенциалу, а нагретое тело охлаждается до температуры окружающей среды. В соответствии с этим же основным законом электроны в системе зон твердого тела стремятся занять самые низкие энергетические уровни. Мы уже установили, что если в зоне меньше электронов, чем требуется для ее заполнения, электроны заполняют уровни зоны снизу вверх до исчерпания их, причем верхняя граница их распределения слегка размыта.

Рассмотрим теперь механизм заполнения энергетических уровней в зонах, с тем чтобы понять, как и почему полупроводники ведут себя характерным для них образом.

В левой части рис. 17-1 показаны близко расположенные друг к другу энергетические уровни в валентной зоне кристалла твердого вещества. Для упрощения картины на ней изображено всего лишь 70 или 80 подобных уровней, тогда как энергетические зоны типичных твердых веществ обычно содержат группы

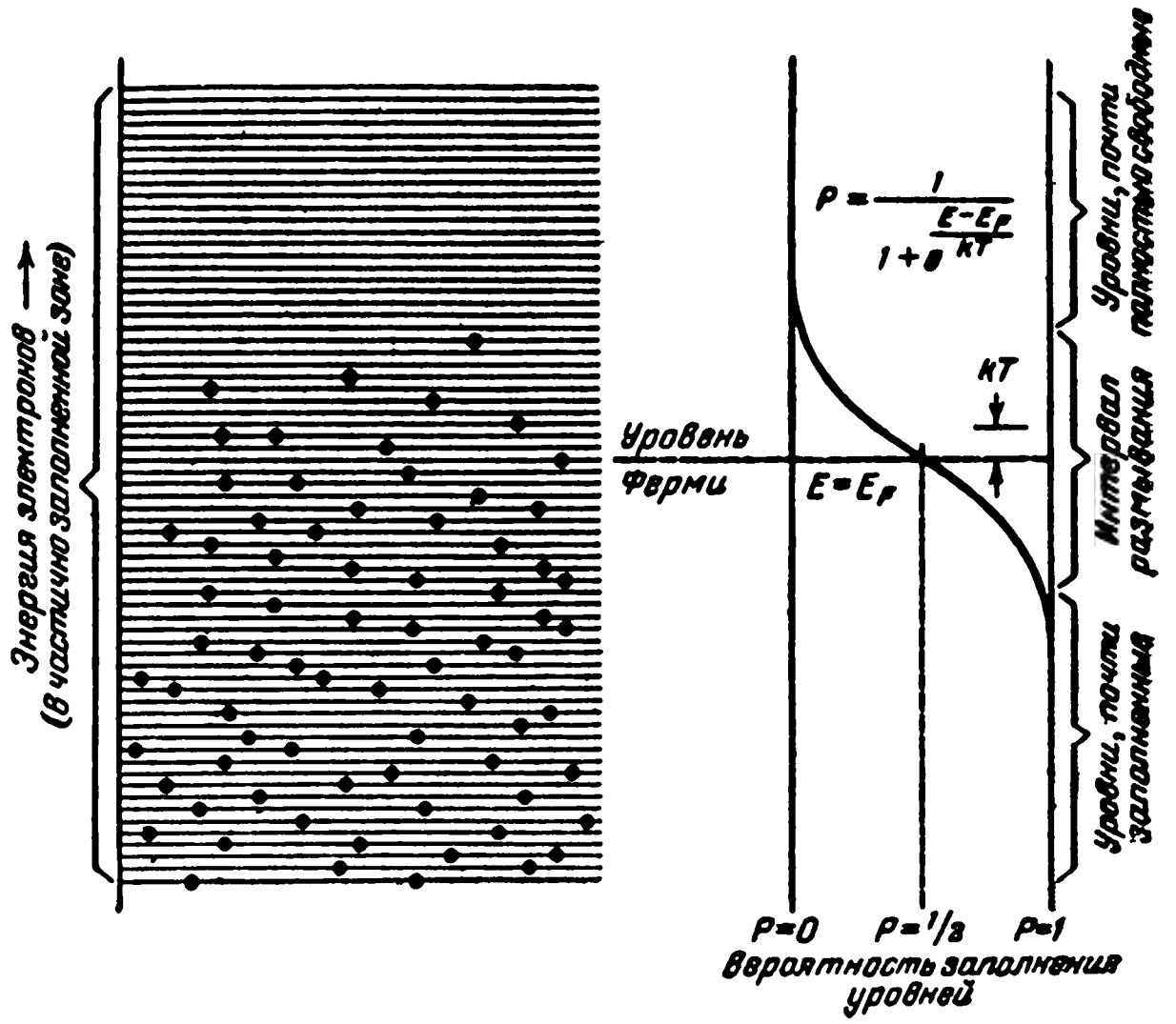


Рис. 17-1. Энергетические уровни электронов в зоне заполняются соответственно функции вероятности, представленной справа.

энергетических уровней числом от 10^{21} до 10^{22} в каждой зоне шириной в несколько электронвольт. В нашем примере предполагается случай вещества, в котором эта зона только частично заполнена электронами. Соответственно это вещество должно быть хорошим металлическим проводником. Энергетические уровни внизу распределения показаны полностью занятыми: каждый уровень занят своими двумя электронами. Рассматривая уровни с большей энергией, можно видеть, что некоторые из них совсем не заняты, либо заняты частично. При еще более высоких энергиях только отдельные уровни оказываются занятыми, а на самом верху зоны уровни почти полностью свободны.

Постепенность перехода от заполненных уровней к свободным при увеличении энергии зависит от температуры: чем выше температура, тем более плавным является переход и тем более размазанной оказывается верхняя граница электронного распределения. Это положение можно сравнить со случаем, когда сосуд с водой сильно встряхнули и на поверхности появились волны, вследствие чего в вершинах волн вода окажется выше среднего уровня поверхности, а ниже этого уровня во впадинах между волнами воды частично недостает. Чем сильнее встряхивают сосуд с водой, тем шире становится по вертикали область перехода от пространства, полностью занятого водой, к пространству, где ее нет совсем. Механическое встряхивание сосуда соответствует действию повышенной температуры в картине электронных уровней; повышение температуры кристалла усиливает колебания атомов его решетки относительно их средних положений, что вызывает возмущение «поверхности» электронного распределения и размытие его границы. При абсолютном нуле температуры тепловые движения внутри кристалла прекращаются, электроны оседают на нижние уровни и у распределения образуется резкая граница перехода от заполненных к свободным уровням. Точно так же у воды в сосуде в состоянии покоя верхняя граница поверхности — резкая: размывание, вызванное рябью, отсутствует.

17-2. СТАТИСТИКА ФЕРМИ

Справа на рис. 17-1 приведен график зависимости вероятности заполнения электронного уровня от энергии этого уровня. Он количественно выражает то, о чем говорилось раньше при обсуждении заполнения уровней. Для нижних уровней вероятность заполнения (двумя электронами) практически полная. Для высоко расположенных уровней (у вершины распределения электронов) вероятность заполнения зависит от энергии уровня согласно графику.

Эта кривая вероятности имеет вид, который можно получить вычислением на основании статистики Ферми:

$$P(E)=\frac{1}{1+e^{(E-E_F)/kT}}; \quad P(V)=\frac{1}{1+e^{e(V-V_F)/kT}}, \quad (17\text{-}1)$$

где E — энергия; V — эквивалентный ей потенциал в вольтах для уровня, вероятность заполнения которого мы хотим исследовать; k — константа Больцмана ($1{,}38 \cdot 10^{-23}$ *дж/град*); e — заряд электрона ($-1{,}6 \times 10^{-19}$ *к*); T — абсолютная температура; E_F — энергия эталонного уровня, на котором вероятность заполнения точно равна $^1/_2$ и относительно которой изображенная кривая вероятности (рис. 17-1) симметрична. Предполагается, что эта энергия соответствует верхней границе электронного распределения при абсолютном нуле температуры, а также средней энергии размытой верхней границы распределения при любой другой температуре T. E_F носит название энергии Ферми или уровня Ферми. Соответствующий ей потенциал V_F называют иногда термодинамическим или электрохимическим потенциалом.

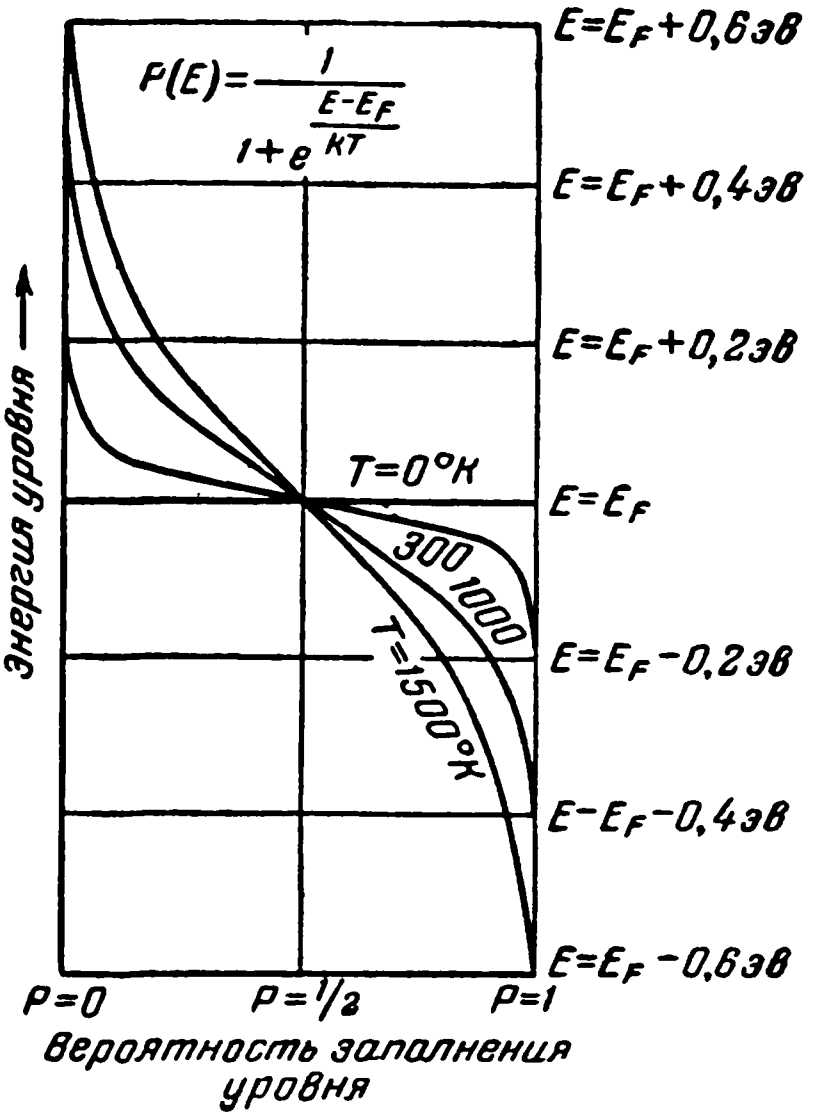


Рис. 17-2. Форма функции Ферми при разных температурах.

На рис. 17-2 показана функция вероятности Ферми, вычерченная для четырех различных температур. Видно, что область энергий вблизи уровня Ферми, в которой происходит большая часть переходов с заполненных уровней на незаполненные, при любой температуре имеет ширину порядка двух или трех kT. При комнатной температуре $kT \approx 0{,}025$ *эв*. Форма этой кривой вероятности не зависит от энергии уровня Ферми, так как P зависит только от разности энергий между уровнем Ферми и исследуемым уровнем. Симметрия кривой вероятности в соседстве с уровнем Ферми показывает, что существует одинаковая вероятность заполнения уровня с энергией, на ΔE большей энергии уровня Фер-

ми, и освобождения уровня с энергией, на ΔE меньшей энергии уровня Ферми.

Концепция вероятности Ферми важна потому, что, применяя ее к энергетическим зонам полупроводников, мы можем вычислить теоретически концентрации электронов и дырок проводимости. Зная это, можно рассчитать удельную проводимость и другие свойства полупроводников и предсказать зависимость их от температуры, содержания примесей и других внешних условий. Функция вероятности Ферми может быть применена не только к вычислению степени заполнения электронных уровней в одной зоне, но также к решению вопросов, связанных с группой зон, энергетических щелей (запрещенных зон) и уровнями донорных и акцепторных примесей.

17-3. СЛУЧАЙ СОБСТВЕННОГО ПОЛУПРОВОДНИКА

В качестве иллюстрации метода и его применений вычислим предполагаемую концентрацию электронов в номинально свободной верхней зоне и концентрацию положительных дырок в номинально заполненной нижней зоне собственного полупроводника, не имеющего ни доноров, ни акцепторов. В левой части рис. 17-3 показана зонная структура такого полупроводника. Ширина энергетической щели между верхним краем валентной зоны и нижним краем зоны проводимости равна E_g. Обозначим плотность электронных энергетических уровней в функции от энергии, измеренной от верхнего края валентной зоны, через $N_1(E)$, а плотность энергетических уровней в функции от энергии, измеренной от нижнего края зоны проводимости, — через $N_2(E)$. Величины $N_1(E)$ и $N_2(E)$ определяются числом энергетических уровней на единицу интервала энергии по оси ординат приведенного выше рисунка. N_1 и N_2 — функции энергии, так как расстояние между энергетическими уровнями меняется при переходе от нижнего края зоны к верхнему. Для пользования статистикой Ферми это, однако, несущественно, так как все энергии следует отсчитывать от некоторого начального (нулевого) энергетического уровня. Это вытекает из того, что в выражение для $P(E)$ входят только разности энергий. Во многих случаях удобно выбрать в качестве нулевой энергию свободного электрона вне кристалла. В дальней-

шем, однако, для простоты мы примем за нуль энергию, совпадающую с верхней границей валентной зоны. Справа на рис. 17-3 показана функция Ферми $P(E)$, соответствующая изображенной картине уровней. При рассмотрении рис. 17-3 сразу выявляется ряд следующих положений:

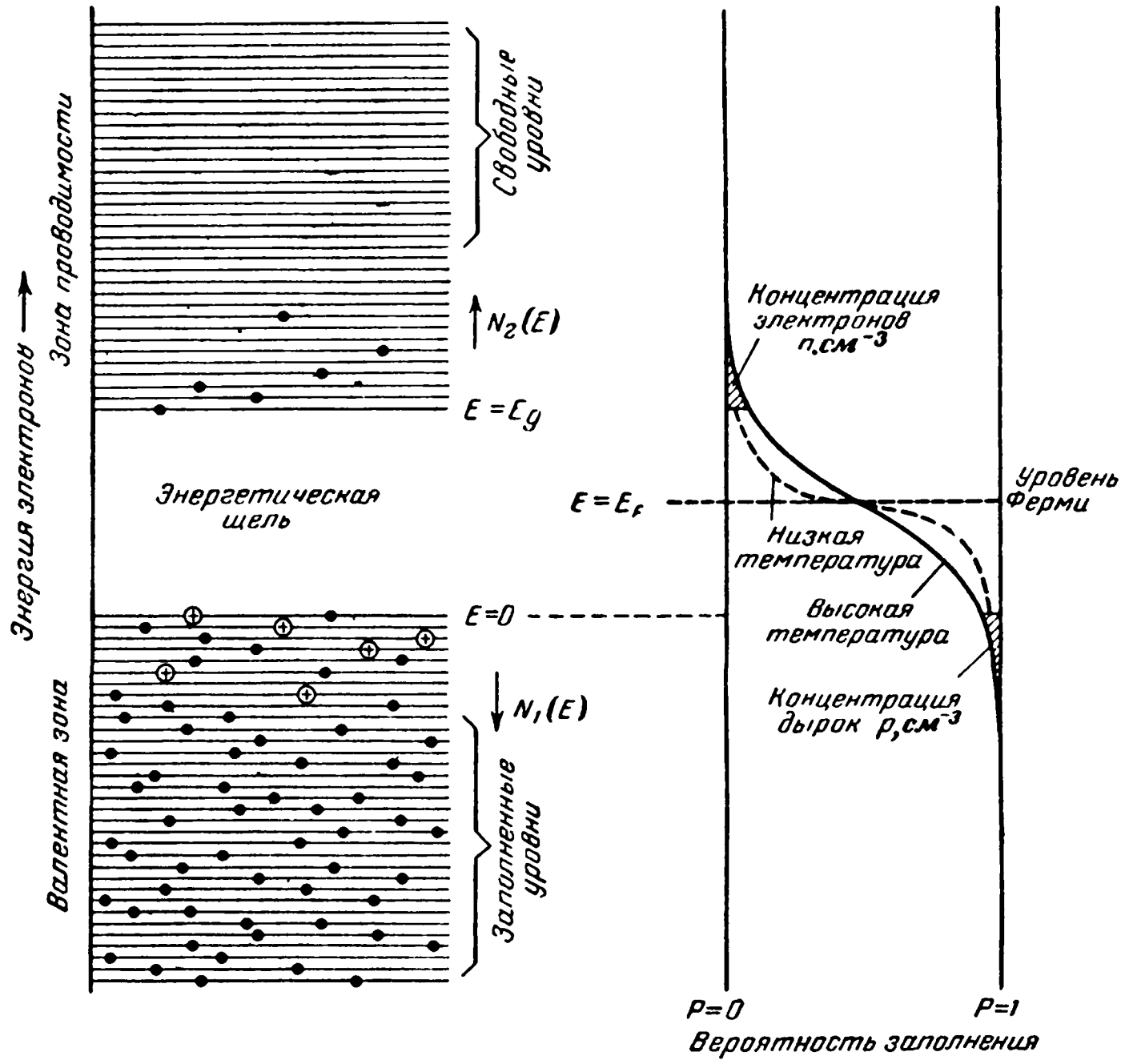


Рис. 17-3. Энергетические зоны собственного полупроводника заполняются соответственно функции Ферми.

1. При всех температурах выше абсолютного нуля в зоне проводимости находится некоторое количество электронов, а в валентной — некоторое число дырок.

2. Количество этих электронов и дырок растет с повышением температуры. Одной из задач настоящего анализа и является выявление характера этого роста.

3. Так как в нашем случае собственного полупроводника нет ни доноров, ни акцепторов, число электронов в кубическом сантиметре полупроводника с энергиями, соответствующими зоне проводимости, всегда равно

числу дырок в кубическом сантиметре полупроводника в валентной зоне. Это равенство есть следствие электрической нейтральности объема полупроводника. Из этого замечания естественно вытекает также, что единственный способ введения электронов в зону проводимости заключается в том, чтобы перевести их из валентной зоны, причем количество оставшихся в валентной зоне дырок всегда точно равно числу переведенных электронов. Эта результирующая нейтральность полупроводника необязательно должна осуществляться в пределах атомных расстояний, но обязательно должна иметь место на достаточно больших участках. Любое значительное отклонение от нейтральности на большом протяжении приведет к возникновению значительного объемного заряда, электрическое поле которого немедленно восстановит эту нейтральность.

4. Уровень Ферми должен находиться где-то в энергетической щели между двумя зонами. Это является следствием того, что в зоне проводимости количество электронов весьма невелико (во много раз меньше, чем число уровней, которые смогли бы принять их), а в валентной зоне столь же мало число вакансий или дырок.

Эти условия соответствуют нижнему и верхнему «хвостам» функции вероятности Ферми, в то время как середина ее (т. е. сам уровень Ферми) должна соответственно находиться где-то между зонами. Кривая на рис. 17-3 справа построена в соответствии с этими соображениями.

Концентрация электронов в зоне проводимости в небольшом интервале энергии dE равна значению функции вероятности Ферми для данной энергии, умноженной на плотность уровней электронов в этой области энергий и на интервал dE области энергии. Общее число n электронов в кубическом сантиметре с энергиями, соответствующими всей верхней зоне, получается в результате интегрирования этого произведения по всем энергиям зоны проводимости от нижнего ее края до верхнего:

$$n = 2 \int_{E_g}^{\text{макс. энергия зоны}} N_2(E)\, P E dE \ см^3. \qquad (17\text{-}2)$$

Коэффициент 2 учитывает возможность занятия каждого электронного уровня двумя электронами. Интеграл уравнения (17-2) может быть точно вычислен. Однако можно получить приближенное решение, применимое к наиболее важным практическим случаям и позволяющее ясно проследить все его физические основы. Заметим, что распределенные в зоне проводимости электроны концентрируются в основном на ее нижних уровнях. Если уровень Ферми отстоит от нижнего уровня этой зоны вниз более чем на несколько kT, то вероятность заполнения уровней, определяемая «хвостом» функции Ферми, спадает экспоненциально с ростом энергии над нижним краем этой зоны. Это означает, что бóльшая часть электронов зоны проводимости должна обладать энергией, лежащей в пределах узкого интервала (всего лишь в несколько раз большего, чем kT), прилегающего к нижнему краю этой зоны. Действительно, интеграл уравнения (17-2) достаточно вычислить для интервала энергий между E_g и (E_g+4kT), чтобы получить 98% всех электронов, находящихся в зоне проводимости. Таким образом, поскольку заняты могут быть только энергетические уровни, находящиеся вблизи нижнего края этой зоны, целесообразно попытаться определить эффективное эквивалентное число таких энергетических уровней, предположив, что все они собраны вместе и совпадают с E_g. Статистическая механика показывает, что для целей расчета все распределенные энергетические уровни зоны проводимости могут быть заменены эквивалентным числом энергетических уровней, совпадающих при энергии $E=E_g$.

Это эквивалентое число есть

$$N_u=\left(\frac{2\pi m^* kT}{h^2}\right)^{3/2} \text{ в } 1\ см^3, \qquad (17\text{-}3)$$

где m^* — эффективная масса электрона; k — постоянная Больцмана, *эрг/град;* h — постоянная Планка ($6{,}54\cdot 10^{-27}$ *эрг·сек*) *. Индекс u относится к верхней зоне.

Таким образом, мы свели распределенные энергетические уровни зоны проводимости к эквивалентному

---

* Уравнение (17-3) не очевидно. Оно получается в результате довольно сложных расчетов, подробности которых можно найти, например, в [Л. 1].

числу N_u уровней, соответствующих одинаковой энергии E_g, причем на каждом из этих уровней может находиться по два электрона.

С помощью подобного же рассуждения можно заменить все распределенные энергетические уровни валентной зоны эквивалентным числом N_L уровней, соответствующих одинаковой энергии $E=0$, на каждом из которых могут находиться два электрона. Следовательно,

$$N_L = \left(\frac{2\pi m^* kT}{h^2}\right)^{3/2} \text{ в } 1\ см^3, \qquad (17\text{-}4)$$

где m^* — эффективная масса положительной дырки, а индекс L относится к валентной зоне.

В дальнейшем мы будем предполагать, что $m^* = m$ — масса покоя свободного электрона. Это не вполне точно, но погрешность, которую мы допускаем при таком предположении, невелика по сравнению с изменениями электрических свойств на порядки величин, происходящими при небольших изменениях температуры, содержания примеси и т. п. Уравнения (17-3) и (17-4) при этом упрощении получают вид:

$$N_u = N_L = \left(\frac{2\pi mkT}{h^2}\right)^{3/2} \text{ в } 1\ см^3.$$

Обозначив не зависящий от температуры множитель в этом выражении через U, получим;

$$N_u = N_L = UT^{3/2} \text{ в } 1\ см^3, \qquad (17\text{-}5)$$

где U — универсальная константа, численно равная $2{,}42 \times 10^{15}\ см^{-3}\ °K^{-3/2}$.

На рис. 17-4 представлена слева та же зонная структура, что и на рис. 17-3. В середине рисунка эта структура представлена для сжатого распределения эквивалентных энергетических уровней, только что полученного нами, а справа — функция вероятности Ферми, которую мы должны применить к ней для численного нахождения концентрации электронов и дырок. Число электронов в 1 $см^3$, имеющих энергии, соответствующие зоне проводимости, выражается через число эквивалентных уровней N_u, умноженное на вероятность заполнения

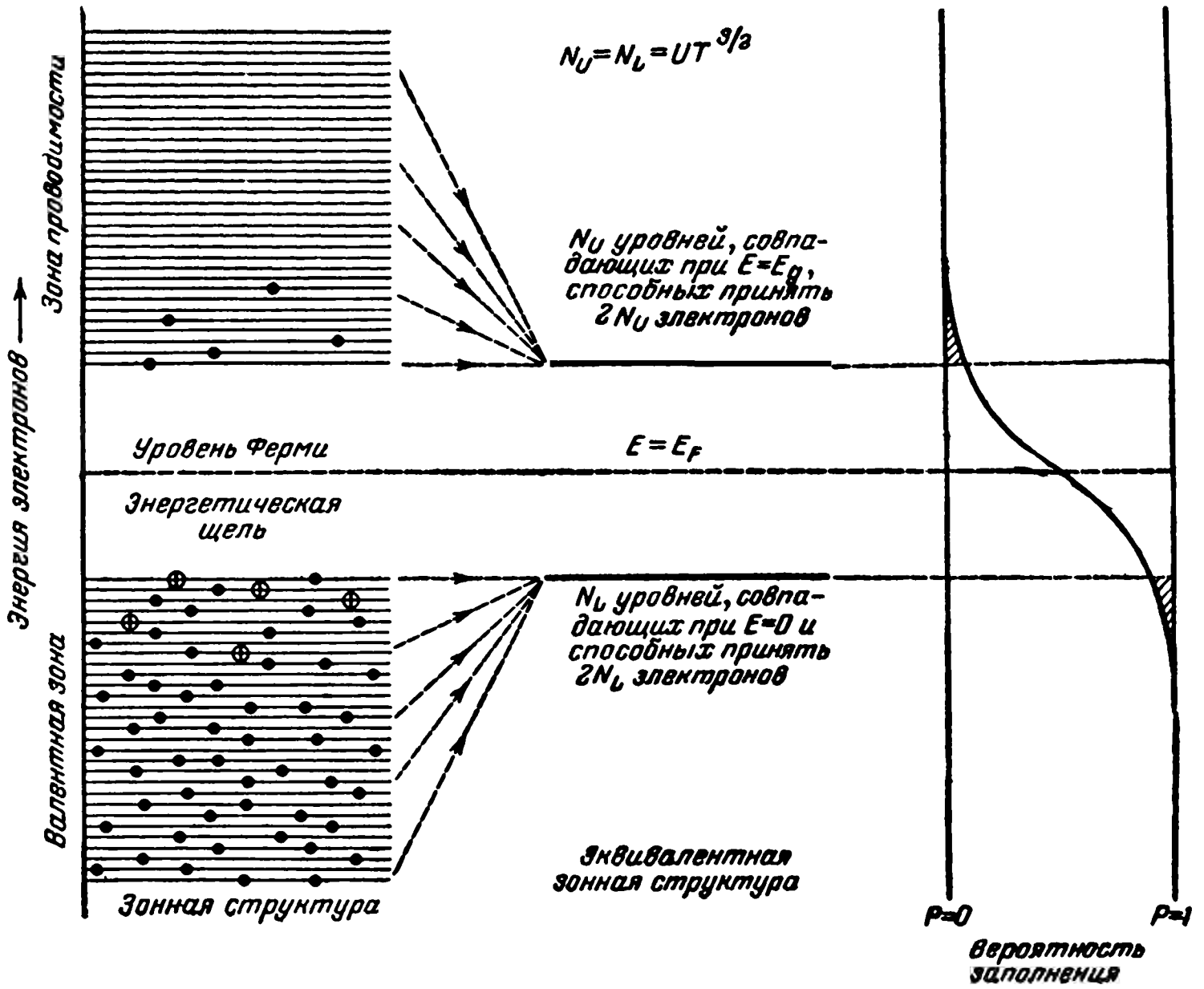


Рис. 17-4. Валентная зона и зона проводимости полупроводника могут быть сведены к двум эквивалентным уровням, которые показаны в средней части диаграммы.

каждого из них и на число электронов, которое каждый уровень может вместить:

$$n = 2N_u \frac{1}{1 + e^{(E_g - E_F)/kT}} \text{ в } 1\ см^3.$$

Подобно этому число дырок в 1 $см^3$ для валентной зоны дается числом электронов в 1 $см^3$, которые могли бы занять все уровни валентной зоны, если бы она была полностью заполнена, минус число электронов, которые в действительности находятся там:

$$p = 2N_L \left(1 - \frac{1}{1 + e^{-E_F/kT}}\right) \text{ в } 1\ см^3.$$

Так как концентрация электронов должна равняться концентрации дырок, то

$$n = p = 2N_u \frac{1}{1 + e^{(E_g - E_F)/kT}} = 2N_L \left(1 - \frac{1}{1 + e^{E_F/kT}}\right),$$

из чего следует, так как $N_u = N_L = UT^{3/2}$, что

$$\left.\begin{aligned} E_F &= E_g/2 \\ \text{и} \qquad n = p &= \frac{2UT^{\frac{3}{2}}}{1 + e^{E_g/2kT}} \end{aligned}\right\}. \qquad (17\text{-}6)$$

В этом уравнении величина E_F выражается теми же единицами энергии, что и E_g. Для всех практических полупроводников E_g должно быть по крайней мере в несколько раз больше kT (в противном случае они не будут полупроводниками). Поэтому, введя в уравнение (17-6) условие, что $E_g > (5—6)kT$, получим:

$$n = p = n_i \approx 2UT^{3/2} e^{-E_g/2kT} \text{ частиц в } 1\ см^3, \qquad (17\text{-}7)$$

где n — концентрация носителей заряда любого типа в собственном полупроводнике. Символ n_i имеет особое применение, которым мы в дальнейшем воспользуемся.

Таким образом, применяя статистику Ферми для собственного полупроводника, мы видим, что в этом случае уровень Ферми соответствует энергии середины энергетической щели и не изменяется при изменении температуры. Концентрации электронов проводимости в зоне проводимости и дырок проводимости в валентной зоне определяются уравнением (17-7). Ввиду того что удельная проводимость полупроводника, содержащего как электроны, так и дырки, определяется уравнением $\sigma = n\mu_n e + p\mu_p e$, мы имеем для удельной проводимости собственного полупроводника выражение

$$\sigma = 2UT^{3/2} e^{-E_g/2kT} (\mu_n + \mu_p) e \ мо/см. \qquad (17\text{-}8)$$

Сравнивая это уравнение с ранее приведенными эмпирическими уравнениями для удельной электропроводности собственного полупроводника, мы убеждаемся, что экспоненциальная зависимость их от температуры полностью согласуется. Кривизна обычно прямых линий зависимости логарифма σ от $1/T$ является результатом

медленного изменения коэффициента $T^{3/2}$ и изменения подвижности в зависимости от температуры. Ясно, что для любого собственного полупроводника ширина энергетической щели E_g или ее эквивалент — напряжение V_g — могут быть получены в первом приближении из наклона графика зависимости [1] $\ln\sigma$ от $1/T$:

$$eU_g = E_g = -2k\frac{d(\ln\sigma)}{d(1/T)}.$$

17-4. ПРИМЕНЕНИЕ ТЕОРИИ К ПОЛУПРОВОДНИКУ *p*-ТИПА

Применим теперь статистический метод к случаю полупроводника *p*-типа, валентная зона которого нормально заполнена, а зона проводимости пуста и имеющего группу локализованных акцепторных уровней.

Как и раньше, обозначим энергию верхнего края валентной зоны через $E=0$, а энергию нижнего края зоны проводимости — через $E=E_g$. Допустим, что локализованные акцепторные уровни имеют энергию E_a и в 1 $см^3$ имеется N_a таких акцепторов. Допустим также во всех последующих рассуждениях, что $N_a \ll N_L$, т. е. что концентрация примеси соответственно никогда не превышает 0,1%. Модель для этого случая показана на диаграмме в левой части рис. 17-5.

Определим далее, где устанавливается уровень Ферми, как на него влияют изменения температуры и какое количество основных носителей заряда (положительных дырок) возникает в зоне проводимости в результате теплового возбуждения электронов с акцепторных уровней при разных температурах. Так же как и в нашем первом примере, энергетические уровни зоны проводимости и валентной зоны могут быть сжаты в тесные группы эквивалентных уровней N_u и N_L при $E=E_g$ и $E=0$ соответственно, как показано на диаграмме в средней части рис. 17-5. Справа на этом же рисунке представлена функция вероятности Ферми, вычерченная рядом с этим изображением эквивалентных сжатых зон. Ее уровень Ферми соответствует некоторой энергии E_F, которую мы и хотим определить.

Относительно положения уровня Ферми при темпе-

---

[1] В этой главе все логарифмы — натуральные.

ратурах ниже области собственных температур полупроводника можно сделать следующие замечания. Электроны попадают на акцепторные уровни при температурах, значительно меньших тех, при которых такое же количество их попадало на уровни зоны проводимости. Сле-

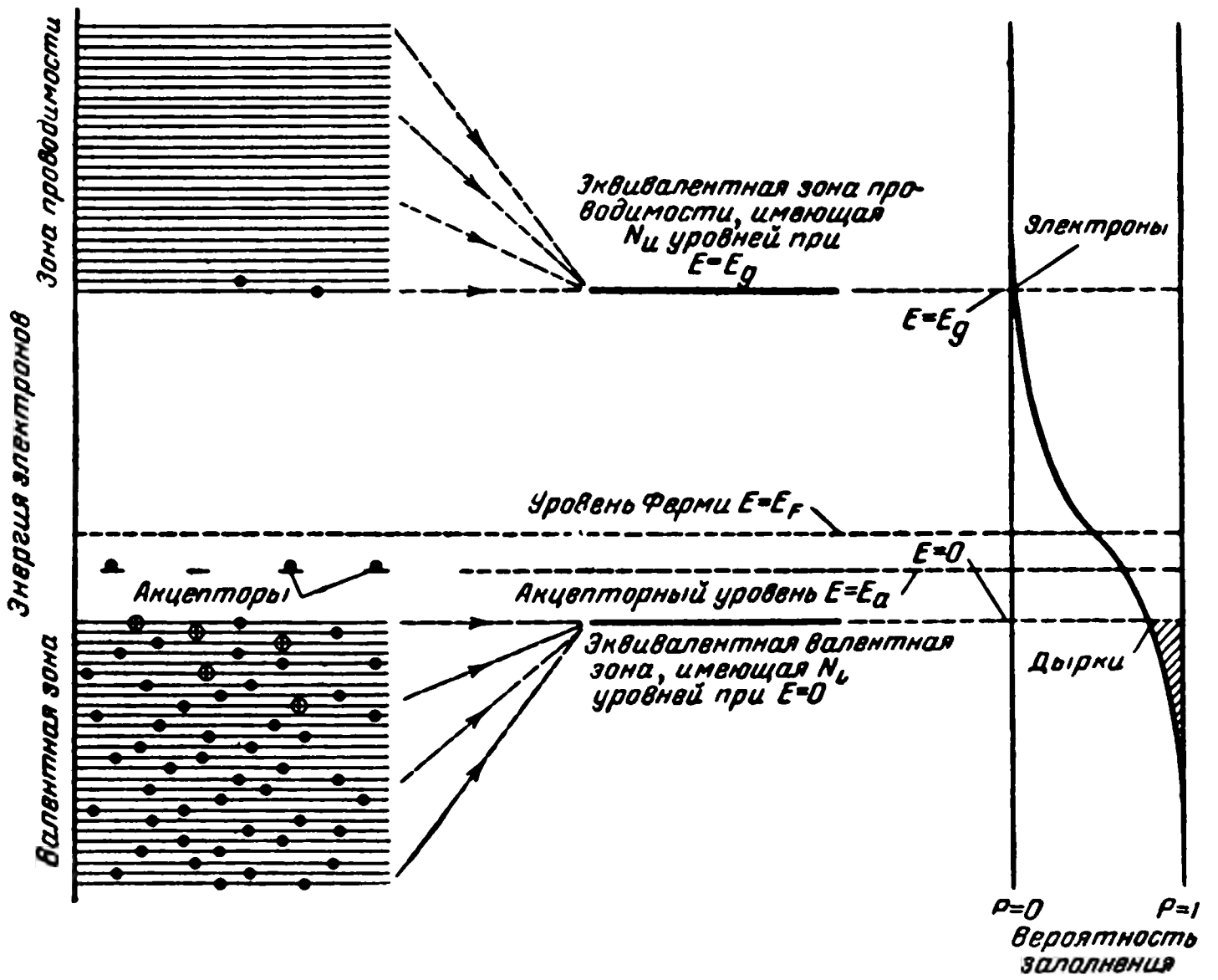


Рис. 17-5. В середине рисунка показана эквивалентная структура сжатых зон-уровней для полупроводника *p*-типа.

довательно, при обычных температурах (меньших области собственных температур) в валентной зоне окажется много больше дырок, чем в зоне проводимости электронов. Уровень Ферми должен в этом случае находиться где-то в нижней половине энергетической щели. Рис. 17-5 начерчен в соответствии с этим заключением. Уравнение, описывающее эту картину, просто устанавливает тот факт, что для сохранения электрической нейтральности полупроводника число дырок в 1 $см^3$ полупроводника в валентной зоне должно равняться числу электронов в 1 $см^3$ на акцепторных уровнях плюс число электронов в 1 $см^3$ в зоне проводимости. Таким образом, применяя функции вероятности Ферми к этим различ-

ным уровням, получим соответствующие концентрации частиц:

$$p = 2N_L\left(1 - \frac{1}{1 + e^{-E_F/kT}}\right) = \underset{\text{дырок в 1 } см^3 \text{ для валентной зоны}}{}$$

$$= \underbrace{\frac{N_a}{1 + e^{(E_a - E_F)/kT}}}_{\text{электронов в 1 } см^3 \text{ на акцепторных уровнях}} + \underbrace{\frac{2N_u}{1 + e^{(E_g - E_F)/kT}}}_{\text{электронов в 1 } см^3 \text{ для зоны проводимости}}. \qquad (17\text{-}9)$$

В константах этого уравнения учитывается, что каждый уровень зоны проводимости и валентной зоны может принять два электрона, а каждый локализованный акцепторный уровень принимает только один электрон. Общее решение этого уравнения для E_F при всех температурах весьма сложно. Можно значительно упростить его, если рассмотреть решения для ряда отдельных температурных областей.

Область 1: температура ниже собственной температуры полупроводника, при которой число электронов в 1 $см^3$ в зоне проводимости пренебрежимо мало по сравнению с числом дырок валентной зоны. Тогда последним членом в уравнении (17-9) можно пренебречь и решение для положения уровня Ферми сводится к

$$E_F \approx kT \ln\left[\frac{N_L}{N_a} - \frac{1}{2} + \right.$$

$$\left. + \frac{1}{2}\left\{\left(1 - \frac{2N_L}{N_a}\right)^2 + \frac{8N_L}{N_a}\, e^{E_a/kT}\right\}^{1/2}\right]. \qquad (17\text{-}10)$$

Это решение можно еще больше упростить для некоторых температур и ограниченных температурных интервалов.

Область 1А: температура достаточно низка, так что экспоненциальный член уравнения (17-10) значительно больше любого другого члена в скобках. Положение уровня Ферми получается из этого упрощенного уравнения (17-10) в виде

$$E_F \approx \frac{kT}{2} \ln \frac{2N_L}{N_a} + \frac{E_a}{2} \approx \frac{kT}{2} \ln \frac{2UT^{3/2}}{N_a} + \frac{E_a}{2}, \qquad (17\text{-}11)$$

а соответствующая концентрация дырок — из уравнений (17-9) и (17-11):

$$p \approx \sqrt{N_a}\sqrt{2UT^{3/2}}\, e^{-E_a/2kT} \text{ в 1 } см^3. \qquad (17\text{-}12)$$

Из уравнения (17-11) видно, что при температуре абсолютного нуля или приближающейся к нему уровень Ферми устанавливается

при энергии $E_a/2$ посредине между верхним краем валентной зоны и энергией акцепторного уровня. По мере роста температуры уровень Ферми сначала слегка понижается из-за отрицательного знака при логарифмическом члене, пока температуры таковы, что $2UT^{3/2}$ меньше, чем N_a, а затем начинает повышаться с ростом температуры; этот рост идет несколько быстрее прямой пропорциональности T вследствие влияния логарифмического коэффициента и продолжается до границ применимости этого приближенного решения.

Зависимость концентрации дырок, пропорциональная согласно уравнению (17-12) корню квадратному из концентрации примесей N_a, связывает концентрацию дырок, а следовательно, и электрическую проводимость, с концентрацией примесей в этой температурной области. Так как содержание примесей в различных образцах одного и того же основного полупроводника может различаться на несколько порядков, легко видеть, что и электрические свойства этих образцов могут сильно различаться при любой температуре в температурной области, где концентрации носителей заряда определяются примесями.

Область 1Б: температура настолько высока, что уровень Ферми совпадает с акцепторным. При этой температуре акцепторные уровни наполовину заполнены, так что $p=N_a/2$ Подставив это значение в уравнение (17-9) и положив $E_F=E_a$, найдем температуру:

$$T=\frac{E_a}{k\ln\left(\frac{4UT^{3/2}}{N_a}-1\right)}\ ^\circ\mathrm{K}. \qquad (17\text{-}13)$$

Уравнение (17-13) можно решить только подстановкой ряда пробных значений температур в уравнение и последующей интерполяцией к тому значению температуры, которое удовлетворяет уравнению.

Область 1В: температура настолько высока, что экспоненциальный множитель в последнем члене уравнения (17-10) становится близким к единице, но не настолько высока, чтобы полупроводник приобрел свойства собственного полупроводника. В этой области положение уровня Ферми определяется уравнением

$$E_F\approx kT\ln\frac{2UT^{3/2}}{N_a}, \qquad (17\text{-}14)$$

а концентрация дырок равна:

$$p\approx\frac{2UT^{3/2}N_a}{2UT^{3/2}+N_a e^{E_a/kT}}\ \text{в 1 } см^3. \qquad (17\text{-}15)$$

Очевидно, что в этой температурной области уровень Ферми продолжает подниматься с повышением температуры, но с бóльшим наклоном, чем в низкотемпературной области 1А. Это происходит ввиду отсутствия коэффициента 2 в знаменателе уравнения (17-14). Зависимость энергии Ферми от температуры продолжает быть более сильной, чем прямая пропорциональность температуре вследствие влияния множителя $\ln T^{3/2}$. Концентрация дырок приближается к N_a

по мере увеличения температуры, показывая, что при этом заполняется все большее число акцепторов К моменту, когда уровень Ферми достигает энергии на $4kT$, более высокой, чем энергия акцепторного уровня, акцепторы оказываются заполненными на 98% и дальнейший рост температуры не влияет практически на заполнение акцепторных уровней и концентрацию дырок. Акцепторы насыщаются, и электропроводность приобретает значения, которые мы раньше отнесли к переходной области температур. Здесь удельная электропроводность почти не зависит от температуры; она зависит от температуры лишь постольку, поскольку зависит от температуры подвижность дырок.

Область 2. Температура настолько высока, что E_F поднялось почти до середины энергетической щели и значительное число электронов возбуждается и переходит из валентной зоны в зону проводимости.

В этом случае необходимо применять уравнение (17-9) в полном виде. Однако при этих условиях акцепторы насыщаются и первый член справа в этом уравнении становится равным N_a. Решение для этой температурной области следующее:

$$E_F = kT \ln \frac{4UT^{3/2}}{N_a + (N_a^2 + 16U^2T^3 e^{-E_g/kT})^{1/2}}, \qquad (17\text{-}16)$$

которое при достаточно высоких температурах дает:

$$E_F = E_g/2$$

и

$$p = N_a + n \approx 2UT^{3/2} e^{-E_g/2kT} \text{ в } 1 \text{ см}^3.$$

Эти результаты совпадают с полученными ранее для собственного полупроводника. Таким образом, статистически подтверждается, что примесный полупроводник p-типа становится собственным при температурах достаточно высоких, чтобы электроны возбуждались через всю энергетическую щель в зону проводимости в таком большом количестве, хотя $p=(N_a+n)$, p и n почти равны. В этих условиях концентрация дырок, обусловленная акцепторным механизмом, перекрывается значительно более высокой концентрацией дырок, обусловленной возбуждением электронов в зоне проводимости.

Как зависит положение уровня Ферми в энергетической щели между валентной зоной и зоной проводимости от температуры и концентрации акцепторов в этих различных областях температур, показано на рис. 17-6. На нем приведены рассчитанные кривые зависимости E_F от T для кремниевых и германиевых образцов p-типа с различными концентрациями акцепторов: 10^{14}, 10^{16} и 10^{18} $см^{-3}$. Постоянные величины E_g и E_a как для кремния, так и для германия взяты из опытных определений [Л. 1]. Кривые для этих двух веществ подобны в том, что они начинаются при $E_F=E_g/2$ около абсолютного

нуля температуры и приближаются асимптотически к $E_F = E_g/2$ при очень высоких температурах. Однако они различаются в том, с какой быстротой изменяется E_F в зависимости от температуры между этими двумя крайними пределами; быстрота этого изменения больше при

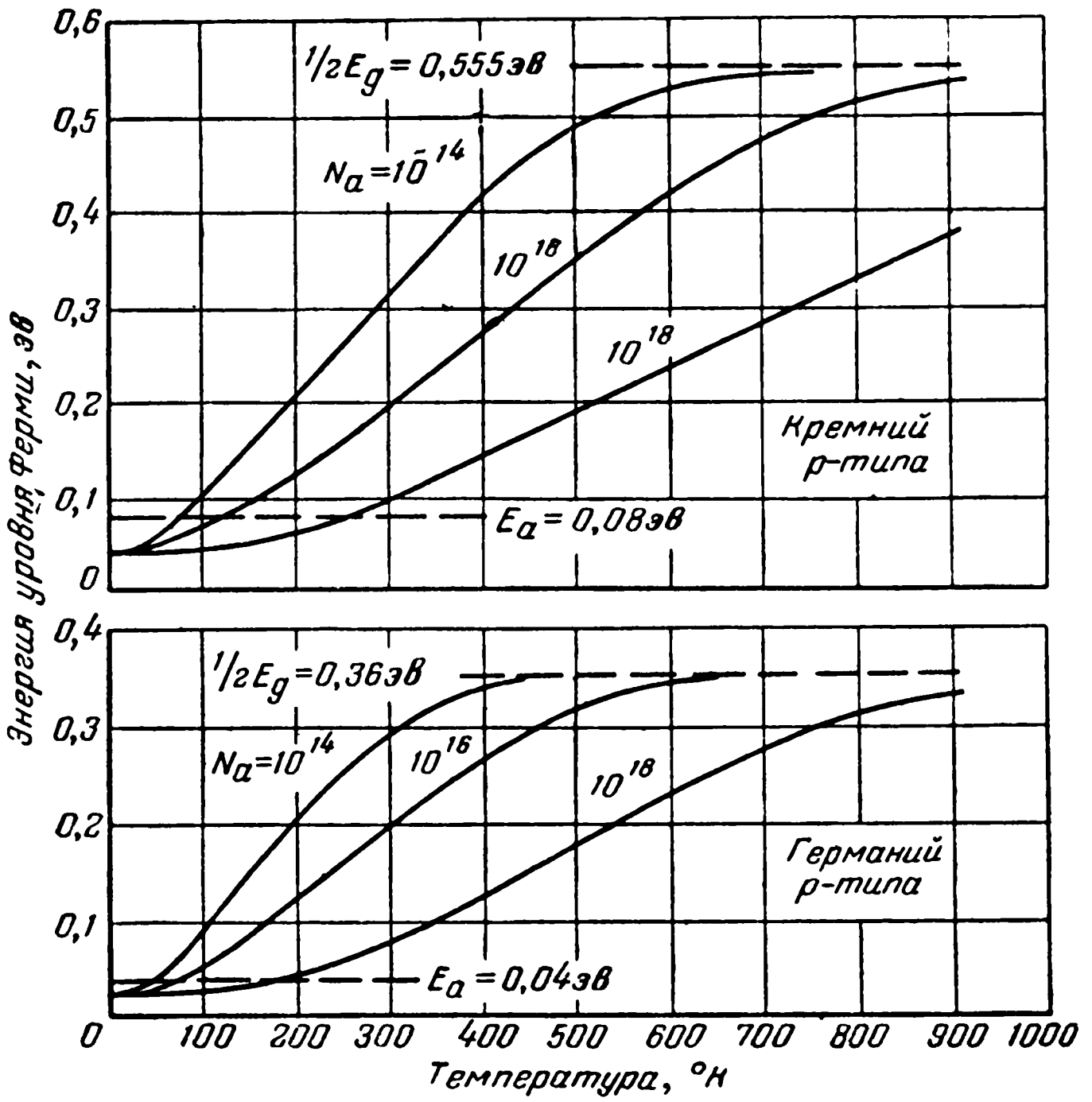


Рис. 17-6 Положение уровней Ферми в образцах кремния и германия p-типа зависит от температуры и концентрации примесей.

малых концентрациях акцепторов. Таким образом, примесный полупроводник с невысокой концентрацией акцепторов превращается в собственный полупроводник при меньшей температуре, чем это имеет место в случае другого образца из такого же полупроводника, но с большей концентрацией примеси.

Все результаты наших обсуждений поведения полупроводника p-типа можно теперь объединить и представить в виде двух графиков на рис.17-7. Верхний график дает зависимость энергии уровня Ферми от температуры

для полупроводника p-типа, а нижний представляет собой логарифм концентрации дырок p в зависимости от обратной величины абсолютной температуры. Линии,

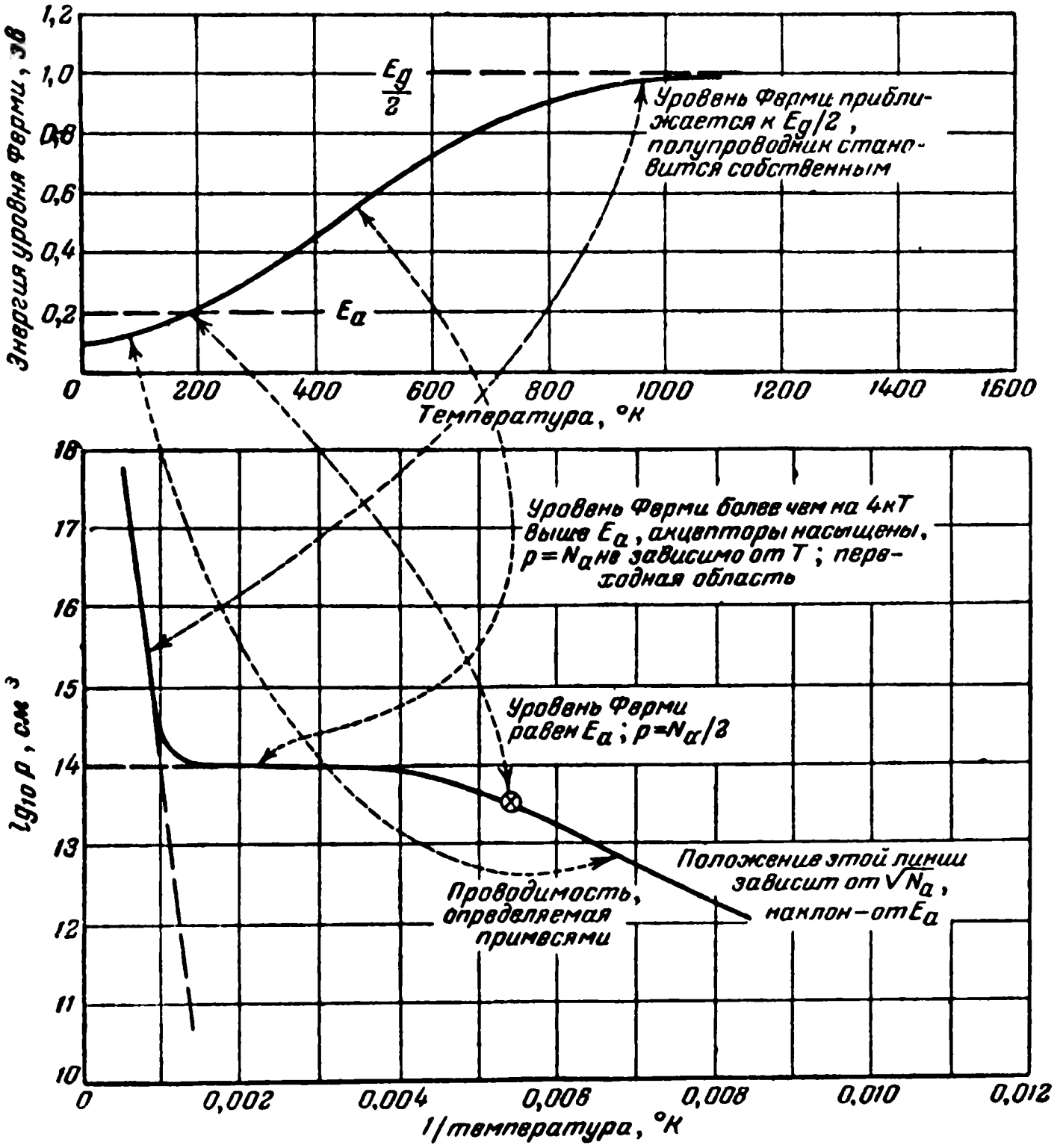


Рис. 17-7. На верхнем графике показано изменение положения уровня Ферми в полупроводнике p-типа в зависимости от температуры, а на нижнем — соответствующие изменения концентрации положительных дырок. Полупроводник имеет $E_g = 2{,}0$ *эв*, $E_a = 0{,}2$ *эв* и $N_a = 10^{14}$ *см*$^{-3}$.

соединяющие обе эти диаграммы при различных температурах, указывают на связь поведения уровня Ферми и концентрации дырок p в различных областях температур. Кривые эти рассчитаны с помощью уравнений, полученных в настоящем параграфе применительно к гипотетическому полупроводнику p-типа, имеющему $E_g = 2$ *эв*; $E_a = 0{,}2$ *эв* и $N_a = 10^{14}$ акцепторам на 1 *см*^{3}.

Обращает на себя внимание сходство полученной кривой зависимости $\log p$ от $1/T$ с экспериментальными кривыми зависимости логарифма удельной электропроводности от $1/T$ (рис. 2-5). Обе кривые имеют следующие общие черты: они стремятся к прямолинейной зависимости с небольшим наклоном при низких температурах, дают почти не зависящую от температуры концентрацию (проводимость) в переходной области температур и имеют прямолинейный участок с большим наклоном при высоких температурах. Отсутствие полного совпадения формы кривой $\log p$ с кривыми $\log \sigma$ обусловлено тем, что удельная электропроводность зависит не только от концентрации носителей заряда, но и от их подвижности, а последняя зависит, как уже указывалось, от температуры. В области собственных температур полупроводника общее число носителей заряда, присутствующих в полупроводнике, вдвое больше числа носителей заряда, указанных на графике для p, так как в зоне проводимости имеется практически столько же электронов, сколько дырок в валентной зоне.

17-5. ПРИМЕНЕНИЕ ТЕОРИИ К ПОЛУПРОВОДНИКУ n-ТИПА

Энергетическая диаграмма для полупроводника n-типа с примесью доноров показана на рис. 17-8. Как и раньше, мы хотим определить, где располагается уровень Ферми и какова концентрация электронов в зоне проводимости при разных температурах. Так же как и раньше, мы упростим задачу, представив зону проводимости и валентную зону N_u и N_L эквивалентными уровнями, совпадающими при энергиях $E=E_g$ и $E=0$ соответственно; предположим, кроме того, что концентрация доноров значительно меньше, чем N_u.

Уравнение, решение которого дает искомые величины, утверждает, что число электронов в зоне проводимости на 1 $см^3$ полупроводника равно числу вакантных уровней доноров в 1 $см^3$ плюс число дырок в 1 $см^3$ полупроводника в валентной зоне.

Таким образом,

$$n=\frac{2N_u}{1+e^{(E_g-E_F)/kT}}=N_d\left(1-\frac{1}{1+e^{(E_d-E_F)/kT}}\right)+$$
$$+2N_L\left(1-\frac{1}{1+e^{-E_F/kT}}\right). \qquad (17\text{-}17)$$

Решение этого уравнения легче всего провести так же, как уравнения (17-9), используя различные приближения в ряде температурных областей. Проделав анализы, аналогичные выполненным в § 17-4, получим следующие результаты для различных температурных областей:

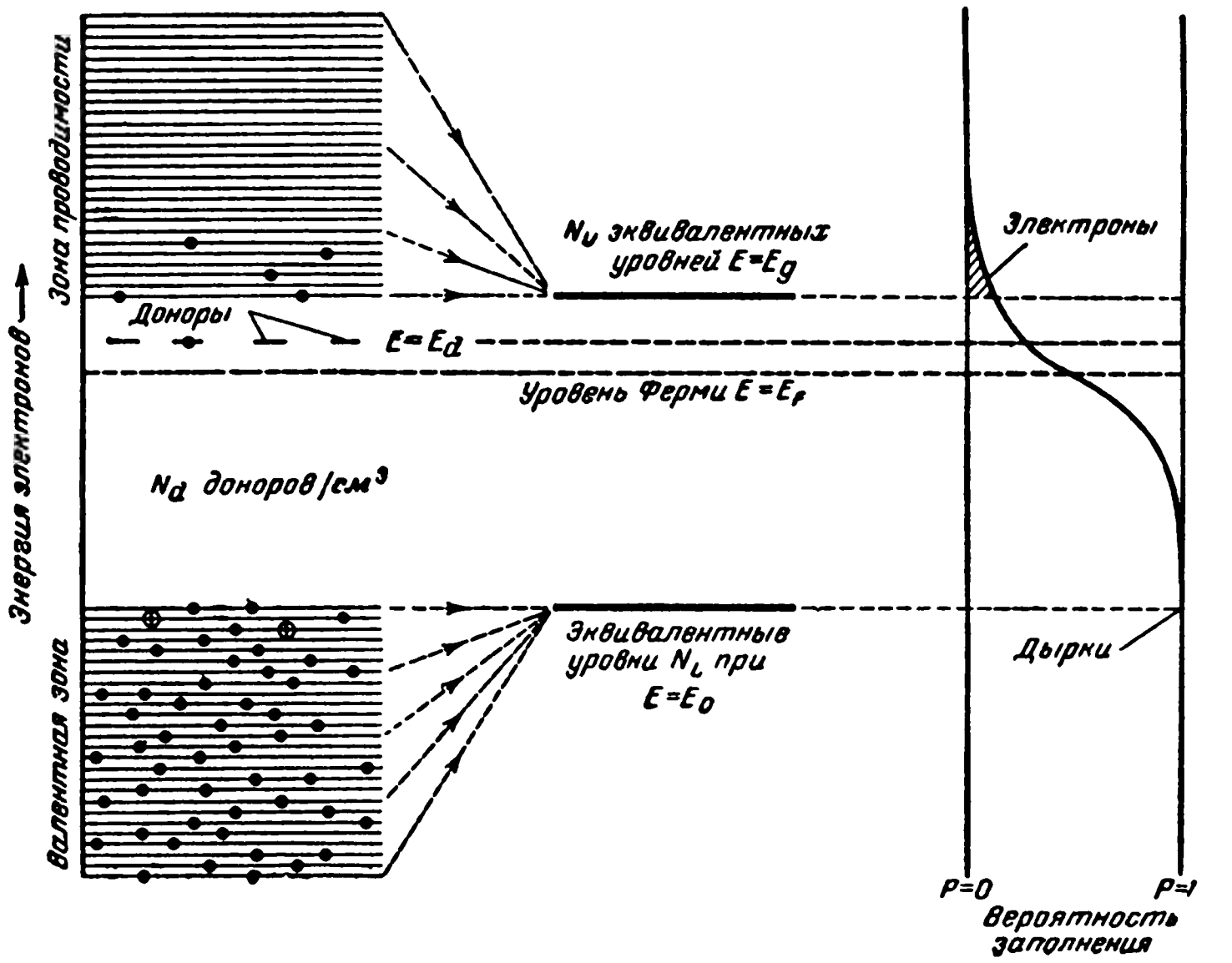


Рис. 17-8. Структура энергетических уровней полупроводника *n*-типа и ее эквивалент в виде дискретных уровней показаны на левой и средней частях рисунка.

Область 1А: при абсолютном нуле и при температуре, близкой к нему, уровень Ферми находится посредине между уровнями доноров и нижним краем зоны проводимости. С повышением температуры от абсолютного нуля уровень Ферми сначала слегка сдвигается вверх, после чего он начинает опускаться в энергетическую щель несколько быстрее, чем пропорционально T. Выражения для E_F и n в этом температурном интервале следующие:

$$E_F = -\frac{kT}{2}\ln\frac{2UT^{3/2}}{N_d}+\frac{E_g+E_d}{2} \qquad (17\text{-}18)$$

и

$$n=\sqrt{N_d}\sqrt{2UT^{3/2}}\,e^{-(E_g+E_d)/2kT} \text{ в 1 } см^3. \qquad (17\text{-}19)$$

Область 1Б: по мере того как уровень Ферми с повышением температуры перемещается вниз, он пересекает уровень энергии доноров. Это происходит при температуре

$$T = \frac{E_g - E_d}{k \ln \left(\frac{4UT^{3/2}}{N_d} - 1 \right)} \text{ ° K.} \qquad (17\text{-}20)$$

При этом условии $n = N_d/2$.

Область 1В: при еще более высоких температурах уровень Ферми сдвигается далее вниз к середине энергетической щели, причем быстрота сдвига все увеличивается Когда энергия уровня Ферми приближается к величине, на $4kT$ меньшей уровня энергии доноров, 98% доноров оказываются свободными и дальнейший рост температуры лишь незначительно влияет на концентрацию n. Удельная электропроводность полупроводника при этом соответствует переходной области. В этой температурной области

$$E_F = E_g - kT \ln \frac{2UT^{3/2}}{N_d} \qquad (17\text{-}21)$$

и

$$n \approx N_d \text{ в } 1 \text{ см}^3. \qquad (17\text{-}22)$$

Область 2: при еще более высоких температурах электроны в результате возбуждения начинают перебрасываться через энергетическую щель все в бо́льших количествах, концентрации n и p продолжают расти, причем их отношение приближается к единице. Полупроводник становится собственным, к которому применимо уравнение (17-8).

Поведение полупроводника n-типа в различных областях температур можно получить и без расчетов, просто по аналогии со случаем уже рассмотренного полупроводника p-типа.

Зависимость концентрации n электронов в полупроводниках n-типа от температуры идентична зависимости концентрации p от температуры для полупроводников p-типа, имеющих такую же величину E_g, $N_a = N_d$ и $E_a = (E_g - E_d)$ для проводника n-типа.

На рис. 17-9 показаны рассчитанные зависимости E_F от температуры для трех образцов кремния n-типа и трех образцов германия n-типа. Все три образца отличаются один от другого только концентрациями примесей доноров. Следует особо отметить симметрию этих кривых и кривых на рис. 17-6 относительно середины энергетической щели.

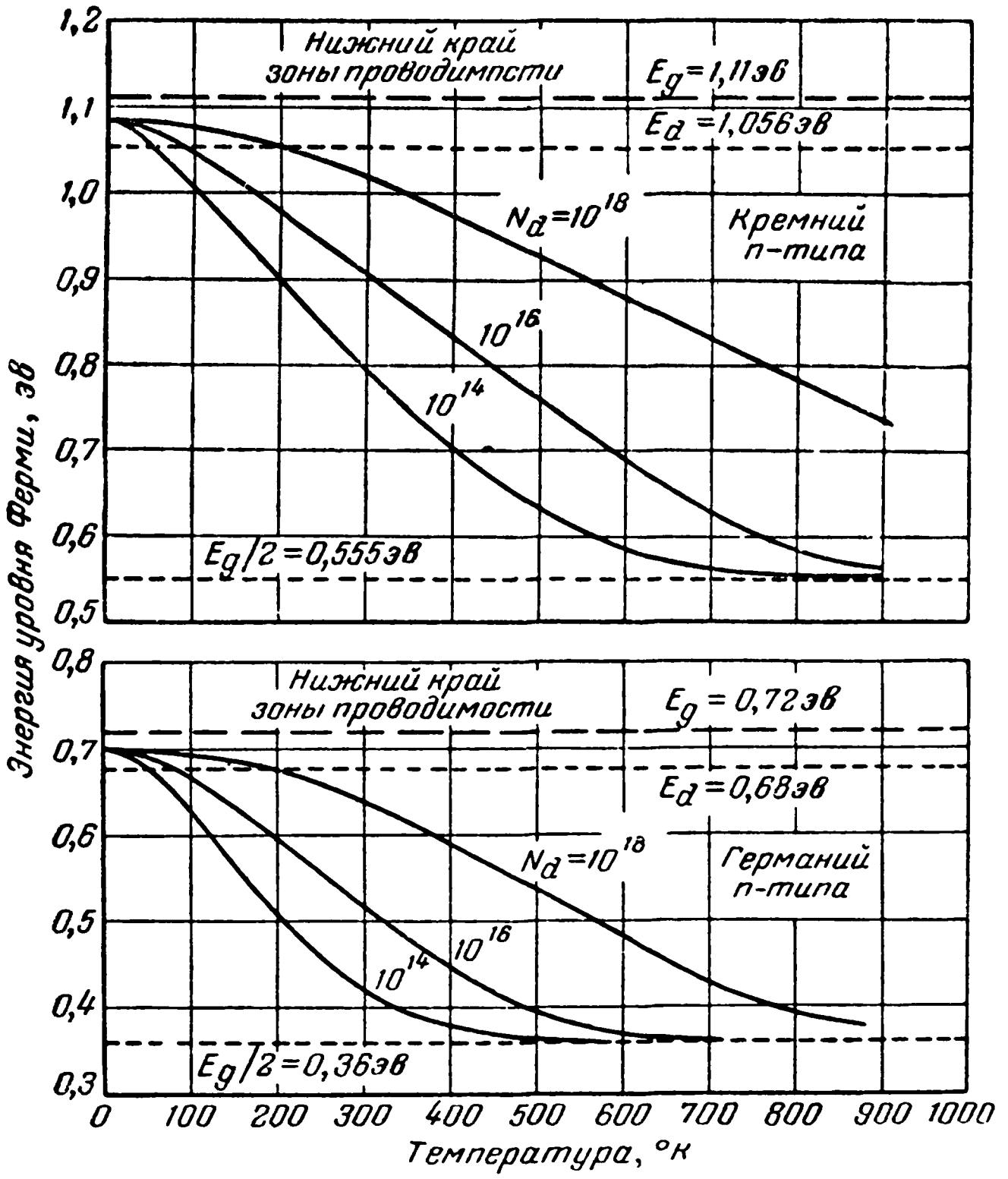


Рис. 17-9 Положение уровней Ферми у образцов германия и кремния *n*-типа в зависимости от температуры и концентрации доноров

17-6. ПРИМЕНЕНИЕ ТЕОРИИ К ПОЛУПРОВОДНИКУ, СОДЕРЖАЩЕМУ КАК ДОНОРЫ, ТАК И АКЦЕПТОРЫ

Это — наиболее общий из всех рассматриваемых случаев. У полупроводников, содержащих как доноры, так акцепторы, электрические свойства определяются в основном тем типом примеси, который численно преобладает. Кремний и германий, в частности, являются примерами такого типа полупроводников. Из них германий может быть очищен с помощью современных технических средств так, что он оказывается собственным полупроводником уже при комнатной температуре. Свойства собственного полупроводника могут быть также прида-

ны и примесному полупроводнику, если добавить в него точно такое же количество примеси противоположного типа для компенсации первоначальной примеси.

Таким образом, можно создать либо чистый собственный полупроводник, либо «компенсированный собственный полупроводник». Однако насыщение образца нейтрализованными примесными атомами сокращает

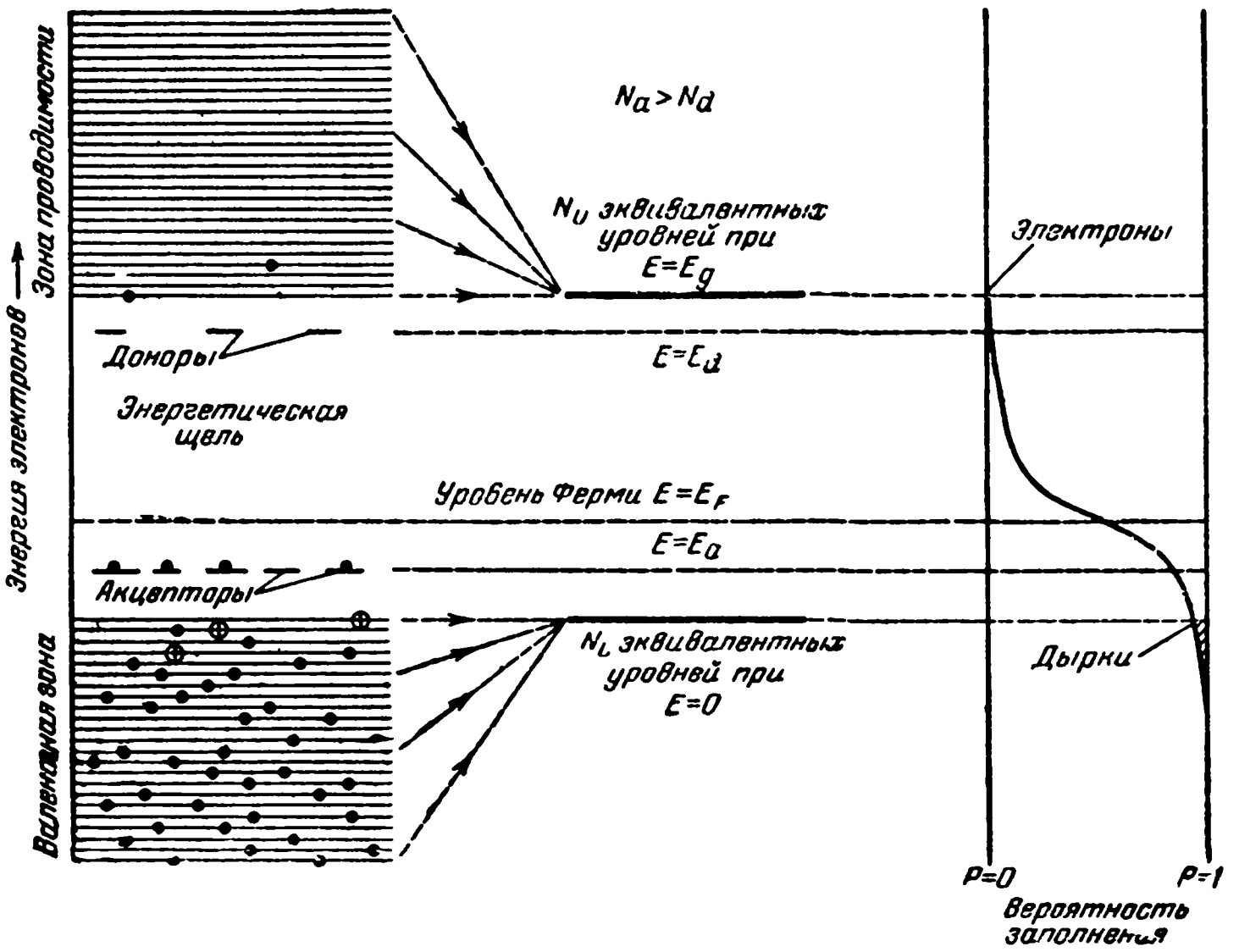


Рис. 17-10 Полная энергетическая схема и эквивалентная энергетическая схема для полупроводника, содержащего как доноры, так и акцепторы, $N_a > N_d$.

время жизни неосновных носителей и уменьшает подвижности обоих типов носителей заряда. Поскольку кремний и германий являются химическими элементами, их состав не может отклоняться от стехиометрического и единственно возможными донорами или акцепторами являются в них посторонние атомы других веществ или рис. 17-10 показана схема энергетических уровней для такого полупроводника.

Основное уравнение для этого случая следует из утверждения, что общий отрицательный заряд должен равняться общему положительному заряду. Таким обра-

зом, концентрация электронов проводимости плюс концентрация связанных электронов на акцепторных уровнях должна равняться концентрации дырок в валентной зоне плюс концентрация освобожденных донорных уровней.

Это выражается следующим уравнением:

$$\frac{2N_u}{1+e^{(E_g-E_F)/kT}}+\frac{N_a}{1+e^{(E_a-E_F)\,kT}}=$$

$$=2N_L\left(1-\frac{1}{1+e^{-E_F/kT}}\right)+$$

$$+N_d\left(1-\frac{1}{1+e^{(E_d-E_F)/kT}}\right) \text{ в } 1\ см^3. \qquad (17\text{-}23)$$

В большинстве встречающихся на практике полупроводников этого типа численно преобладают либо доноры, либо акцепторы. Предположим, например, что преобладают акцепторы, и найдем решение уравнения (17-23) для случая $N_a > N_d$. При температуре абсолютного нуля или близкой к ней

$$E_F = E_a - kT \ln \frac{N_a - N_d}{N_d}\,. \qquad (17\text{-}24)$$

Это показывает, что для данного случая при абсолютном нуле температуры уровень Ферми совпадает с энергией акцепторного уровня. Интересно отметить, что, если $N_d < N_a < 2N_d$, то логарифмический коэффициент имеет положительный знак, а уровень Ферми с ростом температуры *повышается*. Если $N_a > 2N_d$, то логарифмический коэффициент имеет отрицательный знак и уровень Ферми с повышением температуры *понижается*. Если $N_a = 2N_d$, уровень Ферми совпадает с энергией акцепторного уровня в значительном температурном интервале (пока справедливо полученное решение). Концентрация p положительных дырок, соответствующая положению уровня Ферми, отвечающего уравнению (17-24), следующая:

$$p = \frac{N_a - N_d}{N_d}\, 2UT^{3/2} e^{-E_a/kT} \text{ в } 1\ см^3. \qquad (17\text{-}25)$$

Если с ростом температуры уровень Ферми стремится к понижению, то график зависимости E_F от T вскоре меняет свой наклон на обратный и уровень Ферми снова начинает смещаться вверх и пересекает акцепторный уровень при температуре

$$T = \frac{E_a}{k \ln \frac{4UT^{3/2}}{N_a - 2N_d}} \text{ °K.} \qquad (17\text{-}26)$$

При еще более высоких температурах во всех этих случаях уровень Ферми смещается вверх к середине энергетической щели в соответствии с выражением

$$E_F = kT \ln \frac{2UT^{3/2}}{N_a - N_d}. \qquad (17\text{-}27)$$

Это уравнение подобно уравнению (17-14) для простого полупроводника p-типа; разница заключается только в появлении $N_a - N_d$ вместо N_d в знаменателе логарифмического члена. Это подобие свидетельствует о том, что в температурной области, для которой $E_F > E_a$, поведение такого полупроводника можно рассматривать как поведение обычного полупроводника p-типа, в котором эффективная концентрация акцепторов равна $N_a - N_d$. Физически это означает, что N_d электронов с донорных уровней переходят на N_a акцепторных уровней, уменьшая их число на N_d. В результате такой нейтрализации части уровней остается лишь $N_a - N_d$ уровней, эффективно действующих в роли акцепторов. Этот переход электронов на акцепторные уровни является конечным результатом процесса возбуждения электронов из донорных уровней в зону проводимости и последующего перемещения их по кристаллу до встречи со свободными акцепторами, которые и захватывают их. У кремния и германия интервалы температур, в которых возможно такое упрощение, обычно охватывают и комнатную температуру.

При еще более высоких температурах уровень Ферми приближается к середине энергетической щели, доноры и акцепторы перестают определять свойства полупроводника, поведение которого становится таким же, как у собственного полупроводника.

Это же рассуждение можно применить к случаю полупроводника, содержащего в 1 $см^3$ больше доноров,

чем акцепторов. В этом случае при абсолютном нуле температуры уровень Ферми совпадает с уровнем доноров и, по мере того как температура приближается к области собственных температур для данного образца полупроводника, смещается вниз к середине энергетической щели.

Определение концентрации носителей заряда n и p для полупроводников, содержащих примеси обоих типов, выполняется с помощью того же метода, что и в § 17-4 для полупроводников с примесью одного типа.

Рассуждая чисто теоретически, мы, таким образом, получили результаты, которые совпадают с известными фактами об удельной электропроводности полупроводников. Это подтверждает наше первоначальное предположение о том, что описанная картина энергетических уровней верна и что использованный метод пригоден как для предсказания поведения полупроводника, так и для дальнейшего уяснения процессов, происходящих в полупроводнике.

17-7. ПОСТОЯНСТВО ПРОИЗВЕДЕНИЯ pn

Из всего изложенного выше ясно, что при всех температурах, бóльших абсолютного нуля, одновременно в зоне проводимости находятся электроны, а в валентной зоне — дырки. Добавляя к полупроводнику, скажем, примеси донорного типа, можно сильно увеличить число электронов в зоне проводимости. Это приведет к перемещению уровня Ферми в верхнюю половину энергетической щели, что в свою очередь приведет к уменьшению числа положительных дырок в валентной зоне ниже числа, ожидаемого для собственного полупроводника. Действительно, добавление примеси в полупроводник для повышения концентрации основных носителей заряда всегда ведет к понижению концентрации неосновных носителей во столько же раз [1].

---

[1] Это утверждение вполне справедливо, если уровень Ферми отстоит больше чем на (2-3) kT от краев зоны, так что n и p получаются интегрированием с помощью экспоненциальных частей (хвостов) функции распределения Ферми. Это требует, чтобы E_g было не меньше чем по крайней мере (5—6) kT. Если это не имеет места, вещество окажется не полупроводником, а скорее металлом или почти металлом. Для германия $E_g \approx 30\ kT$ при комнатной температуре; для кремния $E_g \approx 45\ kT$

Таким образом, произведение концентрации основных носителей заряда на концентрацию неосновных носителей заряда всегда постоянно для данного полупроводника при данной температуре независимо от того, как его активируют примесью.

Это произведение pn можно легко рассчитать для любого полупроводника, величина E_g которого известна, для чего следует просто возвести в квадрат выражение для n или p собственного полупроводника, даваемого уравнением (17-7). Таким образом, получим:

$$pn = n_i^2 \approx (2UT^{3/2})^2 e^{-E_g/kT} \; см^{-6}. \qquad (17\text{-}28)$$

Для германия $pn \approx 6 \cdot 10^{26}$ $см^{-6}$ при комнатной температуре. Для кремния $pn = 4{,}2 \cdot 10^{21}$ $см^{-6}$ при комнатной температуре. Часто приходится говорить о собственной электропроводности полупроводника. Так как $\sigma = en\mu_n + ep\mu_p$, где μ_n и μ_p — подвижности дырок и электронов соответственно, мы получим:

$$\sigma_i = en_i(\mu_n + \mu_p) \; мо/см. \qquad (17\text{-}29)$$

ЗАДАЧИ

1 На самой глубокой орбите (самом низком энергетическом уровне) атома водорода валентный электрон обладает энергией —14,6 *эв*. На следующей орбите энергия его равна —3,4 *эв*. Предположим, что это единственные уровни водорода. Покажите, что при всех обычных температурах уровень Ферми находится ниже уровня — 14,6 *эв* на ничтожно малую величину.

2 В сосуде содержится моноатомный газ из 10^{20} водородных атомов, о которых говорится в предыдущей задаче. Покажите, что при температуре 3 000° К только примерно 32 атома из всех будут возбуждены.

3. Имеется слиток весом 25 *г* германия *p*-типа, имеющий при комнатной температуре удельное сопротивление 5 *ом · см*. Надо превратить его в материал *n*-типа с удельным сопротивлением при комнатной температуре 1 *ом·см*, расплавив его и добавив небольшое количество сурьмы, после чего рекристаллизовать его в таких условиях, чтобы примеси в получившемся слитке были распределены равномерно. Докажите, что для этого потребуется 0,0023 *мг* сурьмы, если предположить, что каждый атом первоначально имевшейся акцепторной примеси будет точно нейтрализован одним атомом сурьмы. Плотность германия 5,32 $г/см^3$; атомный вес сурьмы 121,8.

4. Кусок германия *p*-типа (рис 17-6) имеет в 1 $см^3$ 10^{14} акцепторов такого типа, что их энергетические уровни лежат на 0,04 *эв* выше верхнего края валентной зоны Найдите положение уровня Ферми при температуре 300° К; пользуясь этой величиной, определите концентрации дырок и электронов проводимости Затем определите удельную электропроводность образца.

5. Что можно сказать качественно о схеме энергетических уровней закиси меди на основании факта отсутствия переходной области?

6. Полупроводник p-типа имеет энергетическую щель шириной 2,0 *эв*, концентрацию акцепторов 10^{16} *см*$^{-3}$ при энергии акцепторных уровней $E_a = 0{,}2$ *эв* Определите с помощью полученных выше выражений положение уровня Ферми в интервале температур от 0 до 2 000° К.

7. Подвижность электронов и дырок в полупроводнике в предыдущей задаче зависит от температуры согласно выражению

$$\mu_n = \mu_p = \frac{T}{0{,}00001T^2 + 1}.$$

Рассчитайте график зависимости $\ln \sigma$ от $1/T$ для этого полупроводника в температурной области от 0 до 2 000° К

ЛИТЕРАТУРА

1. Seitz F., Modern Physics of Solids, McGraw-Hill, New York, 1940.

2. Conwell E. M, Properties of Silicon and Germanium, Proceedings of the Institute of Radio Engineers, 1952, vol. 40, p 1327

3. Адирович Э. И. и др, К теории электрофизических свойств германия, Труды Физического института АН СССР, 1956, 8. 127—177.

ГЛАВА ВОСЕМНАДЦАТАЯ

УРАВНЕНИЕ НЕПРЕРЫВНОСТИ ДЛЯ КОНЦЕНТРАЦИЙ ИЗБЫТОЧНЫХ НЕОСНОВНЫХ НОСИТЕЛЕЙ

18-1. ОБЩИЕ ЗАМЕЧАНИЯ ОТНОСИТЕЛЬНО УРАВНЕНИЙ НЕПРЕРЫВНОСТИ

Уравнения непрерывности встречаются почти в каждом разделе теоретической физики и имеют большое значение вследствие того, что они часто являются отправными пунктами для вывода других выражений, описывающих различные черты поведения всевозможных физических величин или свойств.

Часто уравнение непрерывности является лишь математическим выражением того, что что-то приходящее должно снова уходить. Так, например, уравнение непрерывности для несжимаемых жидкостей просто утверждает, что поток жидкости, втекающей в элемент объема, должен быть равен потоку жидкости, вытекающей

из этого элемента объема, если только в элементе объема не содержится источник или сток жидкости. Таким образом, ρ div $\mathbf{v}=0$, где ρ является плотностью жидкости, а $\mathbf{v}$ — вектор ее скорости. Из этого выражения можно получить уравнение движения, линии тока для различных граничных условий и т. д.

Уравнение Пуассона для статического электрического поля в действительности является уравнением непрерывности. Математически оно выражает, что $E=-4\pi Q$, где E — электростатическое поле, а Q — электрический заряд. На общепонятном языке это уравнение означает, что любой избыток в потоке электростатического поля, исходящего из элемента объема, над потоком, входящим в этот элемент объема, обусловлен источником или стоком его (положительным или отрицательным зарядом), находящимся внутри этого элемента объема.

Уравнение Пуассона является отправной точкой для вывода многих употребительных выражений в электростатике для распределения поля при разнообразных конфигурациях зарядов. Оно является также родоначальником уравнения Лапласа $\nabla^2 V=0$ — основного уравнения теории потенциала для пространства, свободного от зарядов.

Другой пример уравнения непрерывности может быть получен из концепции Максвелла о токе смещения в конденсаторе. Несмотря на «уплотнение» электрического заряда при накоплении его в конденсаторе, ток в цепи, содержащей конденсатор, должен быть везде одинаковым, даже между обкладками конденсатора, где проходит ток смещения $(1/4\pi)\,(\partial E/\partial t)$, который точно равен току проводимости в металлических частях цепи; этот ток сопровождается такими же электрическими и магнитными эффектами, как и ток проводимости. Из этой концепции вытекает одно из основных уравнений теории электромагнитного поля.

18-2. УРАВНЕНИЕ НЕПРЕРЫВНОСТИ ДЛЯ КОНЦЕНТРАЦИЙ ИЗБЫТОЧНЫХ НЕОСНОВНЫХ ЧАСТИЦ В ПОЛУПРОВОДНИКАХ

Подобное же уравнение непрерывности, выражающее некоторые из самых общих идей, можно составить для описания поведения концентрации свойств избыточных неосновных носителей заряда в полупроводнике. Так

как большинство свойств транзисторов, рассмотренных в гл. 10, зависят от введения и переноса избыточных неосновных носителей заряда в различных областях полупроводника, очевидно, что теоретическое исследование работы транзистора может быть произведено с помощью подобного уравнения непрерывности и решения его для различных соответствующих граничных условий.

В этом исследовании для нас представляет особый интерес концентрация неосновных носителей $(p—p_0)$ *, где p_0 — небольшая концентрация неосновных носителей, которая обычно находится в полупроводнике в термодинамическом равновесии с концентрацией основных носителей, тогда как p является местной концентрацией неосновных носителей, которая может быть во много раз больше p_0 из-за того, что вблизи эмитирующего контакта или перехода полупроводник освещен или происходит радиоактивная бомбардировка его, или же какое-либо другое воздействие на полупроводник, кроме обычного теплового возбуждения. Если эти источники удалены или воздействие их прекращено, то p, конечно, начинает быстро спадать, приближаясь к p_0, в результате диффузионного ухода части избыточных неосновных носителей из рассматриваемой области и из-за ограниченного времени жизни той части их, которая остается в этой области и погибает вследствие рекомбинации на месте. Величина $(p—p_0)$ может быть функцией как времени, так и положения внутри полупроводника.

В особых случаях, как, например, в случае модуляции проводимости плоскостного диода, $(p—p_0)$ может даже превысить нормальную концентрацию основных носителей заряда в полупроводнике. Такие крайние случаи, однако, не входят в рамки нашего анализа и нетипичны для нормальной работы транзистора. Мы разберем только те случаи, для которых $(p—p_0)$ остается, скажем, меньше одной десятой нормальной концентрации n_0 основных носителей заряда, но и в этом случае $(p—p_0)$ может быть в тысячи раз больше p_0.

Уравнение непрерывности, которое мы сейчас рассмотрим, констатирует, что быстрота изменения концен-

---

* Мы здесь имеем в виду полупроводник n-типа, и поэтому неосновными носителями заряда в нем являются дырки. Конечно, все аргументы, которые мы здесь получим, могут быть одинаково применены к полупроводникам p-типа.

трации избыточных неосновных носителей внутри элемента объема полупроводника равна разности той скорости, с которой избыточные носители попадают в элемент объема или генерируются в нем, и скорости, с которой избыточные неосновные носители вытекают из этого элемента объема или исчезают в нем из-за рекомбинации или захвата, или другого вида стока.

Есть два способа, которыми избыточные неосновные носители заряда могут попадать в элемент объема или выходить из него, а именно: путем диффузии в направлении градиента концентрации и путем перемещения в электрическом поле. Рассмотрим первый процесс. Величина диффузионного потока избыточных неосновных частиц через 1 $см^2$ поверхности, нормальной к градиенту концентрации, есть $\sim D_p \operatorname{grad}(p-p_0)$, где D — коэффициент диффузии[1] частиц (в данном случае — дырок). Если этот градиент меньше со стороны вытекания из данного элемента объема, чем со стороны втекания внутрь него, то ясно, что в этот элемент входит больше частиц, чем выходит из него, вследствие чего концентрация частиц внутри объема возрастает. Эта быстрота увеличения числа частиц в единице объема определяется выражением

$$D_p \operatorname{div} \operatorname{grad} (p - p_0) = D_p \nabla^2 (p - p_0) = D_p \nabla^2 p,$$

так как p_0 не зависит ни от времени, ни от положения в однородном полупроводнике.

Рассмотрим теперь другой процесс, при котором неосновные носители движутся в электрическом поле E. Ток избыточных неосновных носителей через 1 $см^3$ перпендикулярной потоку поверхности равен $\mu_p (p-p_0) eE$, в то время как поток числа частиц равен $\mu_p (p-p_0) E$. Если концентрация неосновных носителей на стороне выхода частиц из рассматриваемого элемента объема меньше, чем на стороне входа их внутрь, то в объеме станет возрастать концентрация частиц. Быстрота этого возрастания в единице объема равна:

$$-\mu_p E \operatorname{grad} (p - p_0) = -\mu_p (E \operatorname{grad} p),$$

---

[1] D — число частиц, пересекающих за единицу времени единицу площади, перпендикулярной единичному градиенту концентрации (1 частица на 1 $см^3$ на 1 $см$). D имеет размерность $см^2/сек$. Численно для германия при комнатной температуре $D_n = 94$ $см^2/сек$ для неосновных электронов и $D_p = 44$ $см^2/сек$ для неосновных дырок.

где знак минус соответствует положительным дыркам в роли неосновных носителей; электронным неосновным носителям соответствует знак плюс, так как электроны движутся против положительного поля.

Концентрация избыточных неосновных частиц может измениться внутри элемента объема и другим путем — вследствие рекомбинационного исчезновения избыточных частиц внутри объема. Быстрота исчезновения в этом случае равна $(p—p_0)/\tau$ на единицу объема, где τ — среднее время жизни неосновного носителя в окружении массы основных носителей заряда, стремящихся к рекомбинации с ним.

Наконец, есть еще одна возможность возникновения или исчезновения избыточных неосновных носителей заряда внутри рассматриваемого элемента объема. Например, может существовать p-n переход с обратным смещением, который ведет себя, как сток для всех попадающих в него неосновных носителей. Или же объем может быть подвергнут облучению, которое освобождает избыточные неосновные носители заряда на его поверхность или во всем объеме. Допустим, что быстрота возникновения или исчезновения избыточных неосновных частиц в результате такого процесса представлена общим членом g.

В таком случае полное уравнение непрерывности, к которому мы все время стремились, имеет вид:

$$\frac{\partial p}{\partial t} = D_p \nabla^2 p - \mu_p (E \operatorname{grad} p) - \frac{p - p_0}{\tau} + g. \quad (18\text{-}1)$$

Член в левой части этого уравнения — быстрота изменения концентрации избыточных неосновных носителей в единице объема. Первый член с правой стороны— быстрота накопления избыточных неосновных носителей в единице объема, зависящая от разности между скоростью диффузионного поступления частиц и скоростью диффузионного вытекания их. Второй член с правой стороны—быстрота накопления частиц в единице объема, зависящая от неодинаковости числа переносимых электрическим полем частиц вследствие градиента их концентрации. Третий член с правой стороны — быстрота спада концентрации избыточных неосновных носите-

лей в единице объема вследствие их рекомбинации Последний член — быстрота возникновения или исчезновения частиц вследствие наличия других источников или стоков для них.

18-3. РЕШЕНИЯ УРАВНЕНИЯ НЕПРЕРЫВНОСТИ ДЛЯ ЧАСТНЫХ СЛУЧАЕВ

Рассмотрим несколько примеров применения этого уравнения для решения ряда практических вопросов. Большинство случаев, в которых применялось уравнение (18-1), можно считать одномерными в геометрическом смысле; в этом случае уравнение непрерывности принимает следующий вид:

$$\frac{\partial p}{\partial t}=D_p\frac{\partial^2 p}{\partial x^2}-\mu_p E_x\frac{\partial p}{\partial x}-\frac{p-p_0}{\tau}+g. \qquad (18\text{-}2)$$

Пример 1

Однородный кусок германия n-типа подвергается воздействию проникающей радиации, так что концентрация дырок растет равномерно во всем образце. Радиация прекращается при $t=0$. Как изменяется после этого p в зависимости от времени?

Решение

Внутри образца нет электрического поля и градиента концентрации. Кроме того, после прекращения радиации в момент $t=0$ не остается никаких источников и стоков и, следовательно, величина g равна нулю. Уравнение (18-2) принимает при этом вид:

$$\frac{\partial p}{\partial t}=-\frac{p-p_0}{\tau}. \qquad (18\text{-}3)$$

Решение:

$$p-p_0=\text{const}\,e^{-t/\tau}. \qquad (18\text{-}4)$$

Здесь величина константы интегрирования равна первоначальной величине $(p-p_0)$ в момент прекращения радиации. Таким образом, концентрация избыточных дырок, вызванная радиацией, спадает экспоненциально со временем с постоянной времени τ. Для образцов германия различных по чистоте и совершенству решетки кристалла время жизни τ может иметь любую величину от нескольких единиц до нескольких тысяч микросекунд Время жизни неосновных носителей заряда в германиевом транзисторе обычно от 50 до нескольких сотен микросекунд.

Пример 2

Очень короткий (узкий во времени) импульс положительных дырок инъектируется импульсным эмиттером в тонкую нить германия n-типа (рис 18-1 вверху). По обе стороны от эмиттера достаточно места для диффузии облака дырок в направлениях $+x$ и $-x$. Как зависит от времени концентрация избыточных дырок после момента инъекции?

Р е ш е н и е

В уравнении непрерывности для этой задачи член $-\mu_p E_x (\partial p/\partial x)$ равен нулю, так как в нити нет тока [1]

Величина g также равна нулю. Уравнение непрерывности, таким образом, принимает следующий вид:

$$\frac{\partial p}{\partial t} = D_p \frac{\partial^2 p}{\partial x^2} - \frac{p - p_0}{\tau}. \qquad (18\text{-}5)$$

Решение:

$$p - p_0 = \frac{N}{\sqrt{4\pi D_p t}} \exp\left(-\frac{x^2}{4D_p t} - \frac{t}{\tau}\right), \qquad (18\text{-}6)$$

где N — полный введенный первоначально избыточный заряд.

Чтобы это решение было математически точным, протяженность в направлении x введенного облака частиц при $t=0$ должна быть значительно меньше полуширины распределения из (18-6) для любого последующего времени, для которого мы ищем решение Из уравнения (18-6) видно, что облако введенных избыточных неосновных частиц: 1) превращается в результате диффузии в распределение, имеющее форму функции погрешности; 2) спадает из-за рекомбинации по мере того, как оно распространяется, причем оба явления происходят одновременно. На среднем чертеже (рис. 18-1) показан ряд таких решений для последовательных интервалов времени после введения импульса.

Пример 2А

Этот пример во всем совпадает с примером 2, за исключением того только, что в нити поддерживается электрическое поле E

Р е ш е н и е

Член $-\mu_p E_x (\partial p/\partial x)$ должен быть теперь сохранен в уравнении непрерывности. Можно показать, что если решение (18-6) справедливо при отсутствии электрического поля, т. е. имеет форму $(p-p_0)=f(x, t)$, то при наложении электрического поля E решение будет иметь вид:

$$p - p_0 = f(x - \mu_p E_x t, t). \qquad (18\text{-}7)$$

Из уравнения (18-7) видно, что решение примера 2А будет таким же, как и примера 2, за исключением того, что облако заряда избыточных неосновных носителей не только расплывается вследствие диффузии и спадает в результате рекомбинации, но еще и перемещается в направлении x со скоростью $\mu_p E_x$ под действием электрического поля.

В нижней части рис. 18-1 показаны решения для последовательных интервалов времени при наличии такого дополнительного дрейфа.

---

[1] Правда, только что введено облако положительно заряженных дырок Однако пространственный заряд облака втягивает дополнительные электроны в это облако в количестве, необходимом для его нейтрализации Поэтому внутри нити электрического поля практически нет, несмотря на резкие локальные изменения концентрации заряженных частиц.

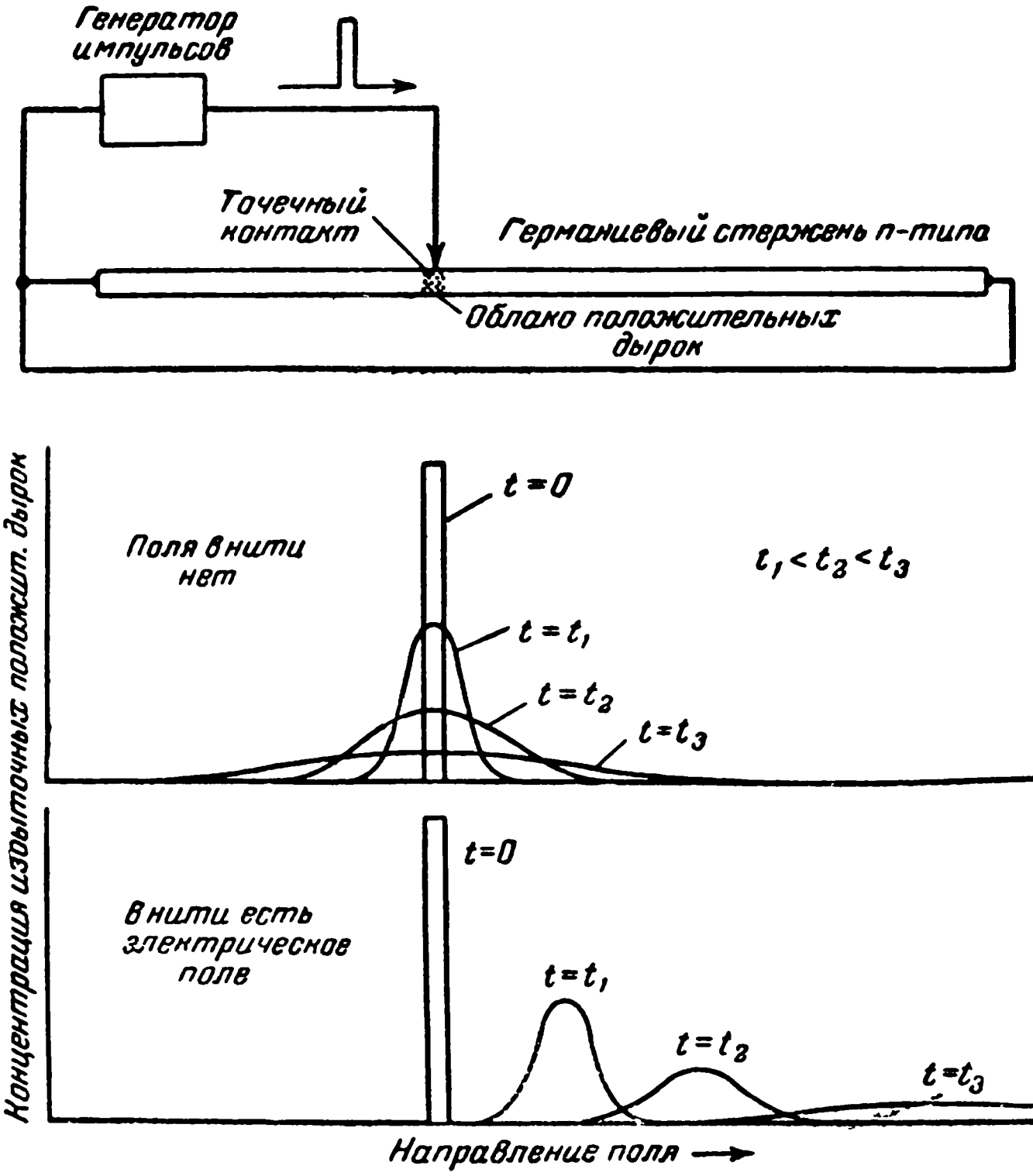


Рис. 18-1. Короткий импульс неосновных носителей, введенный в полупроводник, расплывается вследствие диффузии и спадает из-за рекомбинации. Если имеется электрическое поле (нижняя диаграмма), то облако еще и дрейфует.

Подвижность и время жизни неосновных носителей могут быть измерены практически методом [Л. 1], вытекающим из этого решения уравнения непрерывности. Импульсы неосновных носителей заряда инжектируются в нить с известным электрическим полем. Время, необходимое, чтобы импульс попал под зондовый электрод, расположенный на некотором расстоянии на нити, дает подвижность, а анализ формы проходящего импульса дает время жизни τ. Так как этот анализ легче осуществить на краях ступенчатой функции инжектированного импульса тока неосновных носителей, то практически используются такие функции вместо коротких импульсов, как в примере 2А.

Пример 3

Постоянный ток положительных дырок проходит через p-n переход с прямым смещением в длинный прямоугольного сечения брус полупроводника n-типа. Этот ток достаточен для того, чтобы под-

Держать концентрацию положительных дырок величиной p_1 $см^{-3}$ в плоскости $x=0$ (рис 18-2); эта плоскость расположена достаточно далеко вправо от плоскости p-n перехода, чтобы в ней не действовало электрическое поле и не проникал объемный заряд самого перехода. Каково распределение p в направлении положительных x?

Решение

Мы интуитивно понимаем, что положительные дырки диффундируют вправо от перехода на рис. 18-2, исчезая по пути вследствие рекомбинации, и на некотором расстоянии вправо p спадет до p_0;

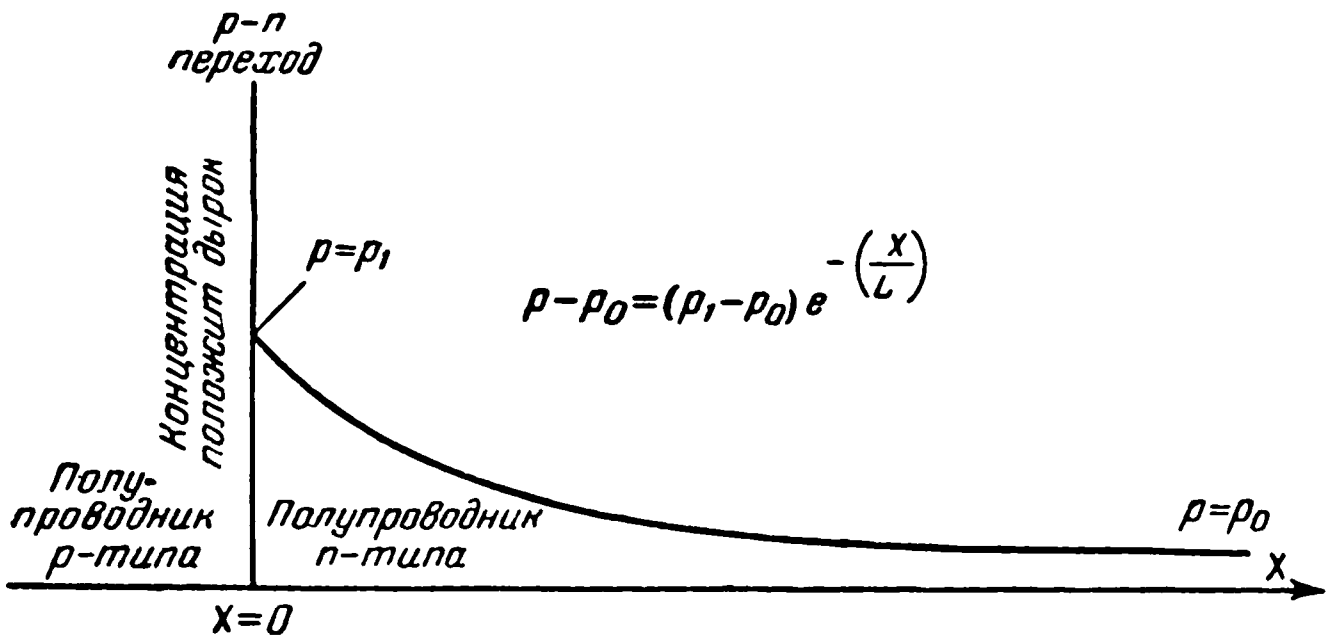


Рис. 18-2. Концентрация избыточных неосновных носителей вблизи p-n перехода, через который проходит прямой ток, спадает экспоненциально с расстоянием от перехода.

бóльшая часть тока за этой плоскостью переносится основными носителями заряда, даже если бóльшая часть тока на переходе и возле него переносится частицами, которые являются неосновными носителями заряда, после того как они прошли через переход.

Так как мы имеем дело со стационарным состоянием, то $\partial p/\partial t=0$. Поскольку во всей n-области бруска с p-n переходом электрическое поле практически отсутствует, член $-\mu_0 E_x(\partial p/\partial x)$ можно считать равным нулю. Правда, имеется небольшое падение напряжения IR, обусловленное током в n-области, но числовым подсчетом можно показать, что электрическое поле, обусловленное этим падением напряжения, незначительно при всех разумных значениях прямого тока. Приложенное внешнее напряжение падает почти полностью на p-n переходе. Так как в n-области нет источников избыточных неосновных носителей заряда, то член g в уравнении (18-2) исчезает и уравнение принимает вид:

$$0=D_p\frac{\partial^2 p}{\partial x}-\frac{p-p_0}{\tau}. \tag{18-8}$$

Решение этого уравнения после введения граничных условий ($p=p_1$ при $x=0$ и $p=p_0$ при $x=\infty$) есть:

$$p-p_0=(p_1-p_0)e^{-x/\sqrt{D_p\tau}}. \tag{18-9}$$

Таким образом, мы видим, что концентрация избыточных неосновных носителей заряда спадает с расстоянием экспоненциально вправо от p-n перехода с постоянной спада $\sqrt{D_p\tau}$. Выражение $\sqrt{D_p\tau}$ часто появляется при анализе поведения полупроводниковых приборов Оно имеет размерность длины, получило название диффузионной длины положительных дырок и обозначается L_p. Понятие диффузионной длины важно, так как, например, если изготовляется транзистор, то его коллекторный электрод нужно помещать в пределах диффузионной длины неосновных носителей заряда в базе; в противном случае потери инъектированных эмиттером

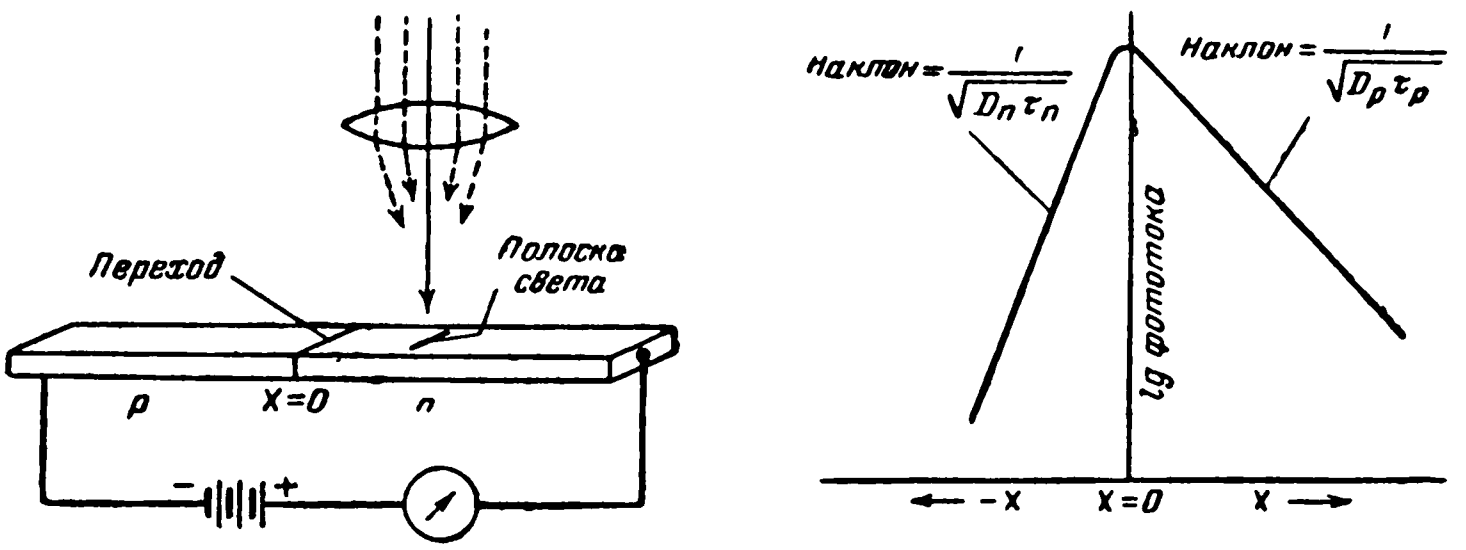


Рис 18-3 Полоска света проектируется на брусок полупроводника с p-n переходом, находящимся над обратным смещением При перемещении полоски света вдоль бруска получается кривая фототока, показанная на правой части рисунка.

частиц вследствие рекомбинации за время перемещения их от эмиттера к коллектору будут чрезмерными. В качестве практического примера можно показать расчетом, что в p-n-p транзисторе, где неосновные носители заряда в базовом слое являются дырками, имеющими время жизни 50 *мксек*, диффузионная длина равна приблизительно 0,05 *см*. Для действительно эффективного собирания зарядов, если желательно, чтобы α транзистора была равна 0,98—0,99, базовый слой не должен быть толще некоторой малой доли диффузионной длины, и необходимой особенностью конструкции обычного транзистора является небольшое расстояние между его эмиттером и коллектором В рассмотренном примере переход со смещением в прямом направлении действовал как постоянный источник новых неосновных носителей заряда. Освещение поверхности полупроводникового образца может также обеспечить непрерывную генерацию неосновных носителей заряда. В частности, если освещение создается в виде узкой полоски света поперек длинной оси тонкой нити полупроводника (рис. 18-3), то в установившемся состоянии концентрация неосновных носителей заряда спадает экспоненциально в обоих направлениях (положительных $+x$ и отрицательных $-x$) от освещенной линии [Л. 2], стремясь к p_0 на расстояниях в несколько диффузионных длин в обоих направлениях. Если на p-n переходе полупроводниковой нити лежит обратное смещение и если полоса света медленно перемещается от одного конца нити к другому, пересекая p-n переход, то обнаруживаемый в результате внутренний фотоэлектрический эффект создает в цепи ток та-

кой же величины, как если бы p-n переход перемещался через создаваемое освещением облако неосновных носителей заряда.

p-n переход с обратным смещением действует как коллектор любых неосновных носителей заряда, диффундирующих в него; поэтому ток через переход является мерой концентрации избыточных неосновных носителей заряда у перехода. Таким образом, если начертить согласно уравнению (18-9) кривую логарифма фотоэлектрического тока цепи в функции положения полоски света, перемещающейся вдоль образца, то получится кривая фоточувствительности, состоящая из двух прямолинейных отрезков с максимумом при абсциссе, где линия света пересекает переход. Кроме того, наклоны этих прямолинейных отрезков, считаемые положительными при движении по абсциссе в любом направлении от перехода, дают прямо величины $1/\sqrt{D_n\tau_n}$ на p-стороны перехода и $1/\sqrt{D_p\tau_p}$ на n-стороне перехода. Так как D_n и D_p численно известны, величины τ_n и τ_p могут быть экспериментально определены с помощью простой аппаратуры и несложного метода обработки.

Мы привели достаточное количество примеров применения уравнения непрерывности для концентрации избыточных неосновных носителей заряда для того, чтобы показать его полезность. Результаты некоторых из этих примеров будут использованы в последующих главах по теории полупроводниковых приборов; по мере необходимости с помощью этого уравнения будут решаться новые задачи.

18-4. СООТНОШЕНИЯ ЭЙНШТЕЙНА

В этом параграфе выведено полезное для дальнейшего соотношение между коэффициентом диффузии D и подвижностью μ частиц для частиц любого типа. Такое соотношение может быть постулировано почти интуитивно, так как беспорядочное поведение частиц определяет как их диффузию, так и статистическую картину подвижности.

Рассмотрим пластину, скажем, p-полупроводника, в которой имеется градиент концентрации акцепторов, т. е. в которой примесь акцепторов является функцией расстояния вдоль этой пластины.

Существование градиента концентрации акцепторов означает также и существование градиенга концентрации положительных дырок, как показано в верхней части рис. 18-4. Так как дырочная концентрация больше у левого конца пластины, то слева направо в единицу времени диффундирует большее число дырок, чем справа налево. Поэтому в правой половине пластины име-

ется небольшой избыток положительного заряда над равновесным. Это перемещение избытка зарядов вызывает появление внутри образца электростатического поля такой величины, что в равновесии обратный ток, обусловленный этим полем, точно уравновешивает диф-

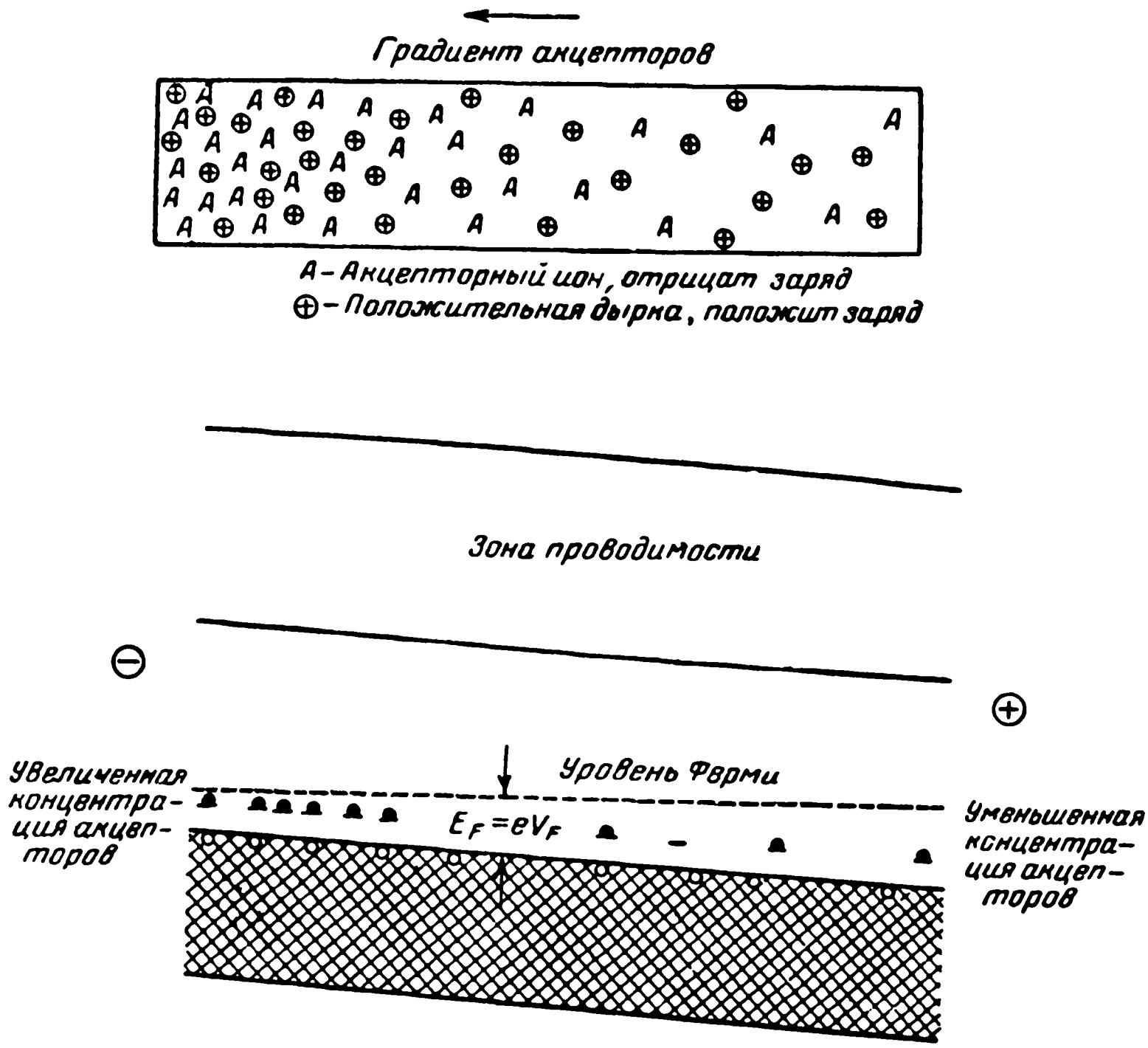


Рис. 18-4 Градиент концентрации примесей в примесном полупроводнике приводит к появлению электростатического поля в направлении этого градиента, что показано наклоном всей системы зон на нижней части рисунка.

фузионный ток. Посмотрим, как выглядит при этом диаграмма энергетических уровней (рис. 18-4, внизу). В последней главе мы видели, что положение уровня Ферми в полупроводнике *p*-типа зависит от концентрации акцепторов N_a, приближаясь к краю валентной зоны при увеличении N_a.

Согласно одному из законов природы, о котором подробнее сказано в следующей главе, в системе, нахо-

дящейся в тепловом равновесии, энергия уровня Ферми постоянна во всей системе. Это означает, что в нашей энергетической диаграмме уровень Ферми должен быть расположен горизонтально, а края энергетических зон наклонены, как показано на рисунке. Величина этого наклона есть мера напряженности электростатического поля, о котором говорилось в последнем абзаце.

Мы можем написать уравнение, которое утверждает, что диффузионный ток положительных дырок, обусловленный градиентом концентрации акцепторов, точно уравновешен обратным током дырок, создаваемым электростатическим полем, который сам является результатом этой диффузии.

Мы имеем:

$$-eD_p \frac{dp}{dx} = -e\mu_p p E_x \ a/см^2. \qquad (18\text{-}10)$$

Из уравнения (17-9)

$$p = 2N_L\left(1 - \frac{1}{1+e^{-(E_F/kT)}}\right) \approx 2N_L e^{-(eV_F/kT)},$$

где последнее приближенное выражение определяется предположением, что $E_F = eV_F > kT$. Тогда

$$\frac{dp}{dx} = \frac{dp}{dV_F}\cdot\frac{dV_F}{dx} \approx -\frac{ep}{kT}\cdot\frac{dV_F}{dx}.$$

Подставив это выражение для dp/dx в уравнение (18-10) и замечая, что $dV_F/dx = -E_x$ — напряженность электростатического поля, мы получаем:

$$\left.\begin{aligned} &D_p \frac{e}{kT} \approx \mu_p, \text{ или } D_p \approx \frac{kT}{e}\mu_p; \\ &\text{таким же образом} \\ &D_n \approx \frac{kT}{e}\mu_n. \end{aligned}\right\} \qquad (18\text{-}11)$$

Эти выражения известны как соотношения Эйнштейна.

Изложенное в этом параграфе приводит, таким образом, к двум важным заключениям. Одно из них за-

ключается в том, что где бы в полупроводнике ни возник градиент концентрации доноров или акцепторов, там обязательно возникает и соответствующее электростатическое поле. Это положение развито в следующей главе и распространено на случай весьма больших градиентов, возникающих внутри барьеров *p-n* переходов. Второе заключение — это соотношение Эйнштейна, выражаемое уравнениями (18-11), которое также будет использовано в дальнейшем.

ЗАДАЧИ

1. Докажите, что уравнение (18-6) является решением уравнения (18-5) и что уравнение (18-7) есть решение уравнения непрерывности для случая наложенного электрического поля.

2. В случае примера 1 этой главы предположите, что интенсивность проникающей радиации изменяется синусоидально в зависимости от времени (от нуля до $I_{\text{макс}}$) согласно выражению $I = \frac{1}{2} I_{\text{макс}}(1+\sin \omega t)$. Быстрота генерации (фотоэлектрического освобождения) новых дырочно-электронных пар в K раз больше мгновенной интенсивности. Составьте и решите уравнение непрерывности, соответствующее этому случаю, показав, что в установившемся состоянии концентрация избыточных дырок равна:

$$p - p_0 = \frac{K}{2} I_{\text{макс}} \tau \left(1 + \frac{\sin \omega t - \omega\tau \cos \omega t}{1 + \omega^2 \tau^2}\right).$$

Вычертите амплитуду этих установившихся колебаний в зависимости от частоты флуктуаций света.

3. Пластина *n*-полупроводника, имеющая параллельные стороны и толщину w в направлении x, подвергается облучению проникающей радиацией, которая генерирует дырочно-электронные пары с одинаковой быстротой g в секунду во всем объеме. Время жизни дырок во внутренней части пластины τ *сек*, но на поверхности быстрота рекомбинации их столь велика, что концентрации дырок на поверхностях все время равны равновесной концентрации p_0. Докажите, что в установившемся состоянии концентрация избыточных дырок в пластине равна:

$$p - p_0 = g\tau \left\{ \left(e^{-w/L_p} - 1 \right) \operatorname{csch} \left(\frac{w}{L_p} \right) \times \right.$$

$$\left. \times \sin \left(\frac{x}{L_p} \right) - \left(e^{-x/L_p} - 1 \right) \right\}.$$

Вычертите эту концентрацию по толщине пластины в области от $0{,}1L_p$ до $10L_p$.

ЛИТЕРАТУРА

1. Haynes J. R., Shockley W., Mobility and Life of Injected Holes and Electrons in Germanium, Physical Review, 1951, vol. 81, p. 835.
2. Goucher F. S. et al., Theory and Experiment for a Germanium p-n Junction, Physical Review, 1951, vol. 81, p. 637.
3. Болтакс Б. И., Диффузия в полупроводниках, Физматгиз, 1961.
4. Губанов А. И., Теория выпрямляющего действия полупроводников, Гостехиздат, 1956.

ГЛАВА ДЕВЯТНАДЦАТАЯ

ТЕОРИЯ *p-n* ПЕРЕХОДА

19-1. *p-n* ПЕРЕХОД В РАВНОВЕСНОМ СОСТОЯНИИ

Мы теперь подготовлены к рассмотрению во всех подробностях физики выпрямляющих свойств *p-n* переходов, чем и займемся, уделив особое внимание использованию их в качестве эмиттеров с прямым смещением и коллекторов с обратным смещением. Несмотря на то, что в первых выпрямителях использовались контакты металл — полупроводник, проще понять физику действия выпрямителя с *p-n* переходом и математически описать его работу. Поэтому рассмотрение контакта металл — полупроводник мы отложим до следующей главы.

Рассмотрение *p-n* переходов мы начнем с обсуждения их свойств, основываясь на зонной структуре полупроводников *p*-типа и *n*-типа, и прежде всего остановимся на природе тока, проходящего через переход, когда прикладывается напряжение смещения. Описав, таким образом, что и почему происходит, мы перейдем к количественному выражению этих закономерностей. В частности, будет выведено выражение для тока, проходящего через переход при прямом или обратном смещении. Будут исследованы условия, при которых прямой ток, проходящий через такой переход, создается главным образом неосновными носителями с одной стороны от перехода, что приводит к высокой эффективности эмиссии неосновных носителей, если переход используется как эмиттер плоскостного транзистора. Кро-

ме того, выясним, как емкость перехода зависит от свойств прилегающего к нему вещества.

На рис. 19-1 схематически показаны энергетические зоны для полупроводников; слева — имеющего некоторую концентрацию акцепторных примесей, превращающих его в полупроводник p-типа, а справа — растворенные донорные примеси, превращающие его в полупроводник n-типа. Предположим сначала, что $N_a \approx N_d$, слу-

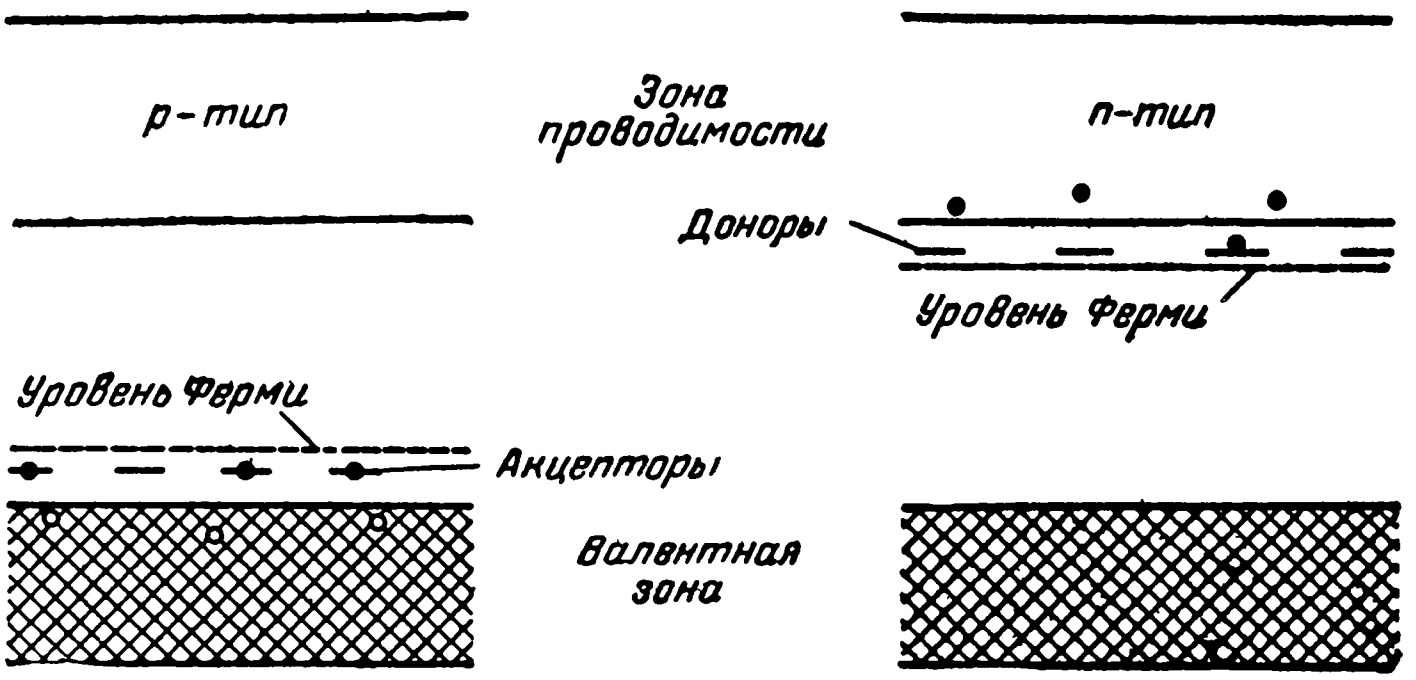


Рис 19-1. У p- и n-типов образцов одного и того же полупроводника уровни Ферми имеют разные энергии.

чай же когда они не равны, рассмотрим позже. Далее, предположим для простоты, что полупроводник p-типа не содержит доноров, а полупроводник n-типа не содержит акцепторов. Заметим, что уровни Ферми в этих двух образцах расположены различно относительно границ зоны.

Предположим теперь, что эти два образца приведены в соприкосновение. Чтобы правильно представить себе, что происходит в этом случае, вообразим, что никаких осложнений, обусловленных возможными поверхностными пленками окиси, или других загрязнений нет, что контакт происходит вдоль всей контактирующей поверхности и что кристаллические решетки двух образцов являются идеальным продолжением друг друга без каких-либо искажений в месте контакта.

Совершенно очевидно, что эти требования неосуществимы для двух отдельных образцов полупроводника, приведенных в соприкосновение. Единственными способами получения таких контактов в настоящее время

являются следующие: выращивание переходов в монокристаллических образцах, вплавление или диффузия примесей одного типа в образцы первоначально противоположного типа или другие процессы, которые создают переходы *внутри целых кусков полупроводника и где угодно, кроме их поверхностей.* Однако для простоты можно предположить, что в соприкосновение приводятся два отдельных образца полупроводника и что при контакте все упомянутые выше требования удовлетворяются. Что же происходит в этом случае на полученном таким образом *p-n* переходе?

Поскольку электронов проводимости больше в зоне возбуждения *n*-области, чем в зоне возбуждения *p*-области, в таком контакте будет наблюдаться диффузионный ток электронов справа налево через переход. Подобным же образом, поскольку положительных дырок больше в *p*-области, чем в *n*-области, дырки будут перемещаться в валентной зоне слева направо через переход. Эти неуравновешенные токи в очень короткий промежуток времени (порядка нескольких миллимикросекунд) приводят к тому, что обе области заряжаются одна относительно другой. Дырочная область приобретает избыток отрицательных зарядов, а электронная область приобретает избыток положительных зарядов. Таким образом, на переходе создается разность потенциалов, причем в таком направлении, чтобы противодействовать дальнейшему неравновесному току. Устанавливается новое равновесие, при котором эта разность потенциалов имеет как раз достаточную величину, чтобы воспрепятствовать стремлению большинства основных носителей заряда проходить через переход под действием градиентов концентрации таких носителей внутри перехода и вблизи него.

На рис. 19-2 изображена энергетическая схема, соответствующая этому окончательному равновесию. Разность потенциалов на переходе приводит к смещению энергетических уровней *n*-области вниз, а уровней *p*-области — вверх. Эти смещения таковы, что уровни Ферми двух областей теперь совпадают и представляются на диаграмме в виде одной горизонтальной линии. Это совпадение является частным примером общего закона, гласящего, *что в системе при тепловом равновесии энергия, соответствующая уровню Ферми, постоянна во всей*

системе. Точно так же, если два сосуда с водой соединены трубкой так, что они образуют одну систему, поверхность воды установится на одном уровне и будет оставаться на нем, пока система не будет нарушена.

Заметим, что, как следует из рис. 19-2, если уровни Ферми двух областей совпадают по энергии, границы зон образуют на переходе ступеньки. Высота этих ступенек V_0 равна контактной разности потенциалов. Числен-

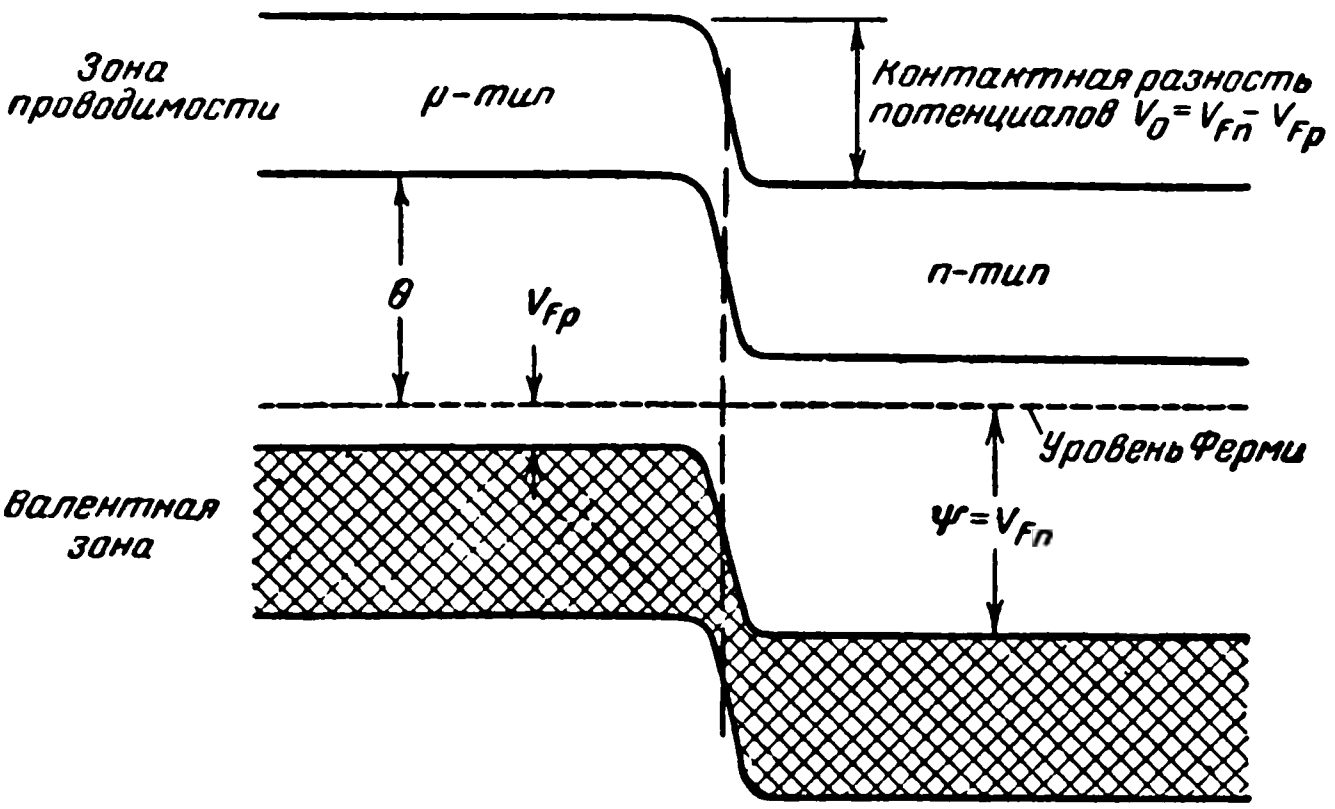


Рис. 19-2. Энергетическая диаграмма для случая контакта двух образцов, схемы которых показаны на рис. 19-1. Энергия уровня Ферми постоянна.

но она равна $(V_{Fn}—V_{Fp})$, где V_F измеряется от границы нижней зоны в каждой области. Наклон ступеньки на переходе и вблизи него равен величине электростатического поля, обусловленного контактной разностью потенциалов. Напряженность этого поля в местах, где оно максимально, может составлять тысячи или десятки тысяч вольт на сантиметр и зависит от активирования прилегающих участков полупроводника. Ширина области, в пределах которой происходит это падение потенциала и возникает поле, обычно порядка от 10^{-4} до 10^{-6} *см*. Эту узкую область резко повышенного градиента напряжения мы называли ранее «барьер».

При первом знакомстве с контактной разностью потенциалов может возникнуть ошибочное представление о сильном «встроенном» электрическом поле, которое поддерживается внутри полупроводника. В этом случае

первой мыслью будет попытаться использовать эту самоподдерживающуюся разность потенциалов как источник э. д. с. путем присоединения электродов к комбинации *p-n* переходов, которые будут в данном случае использоваться как источники мощности. К сожалению, этого не получается! При присоединении электродов неизбежно возникают другие контактные разности потенциалов и общая э. д. с. цепи неизменно равна нулю до тех пор, *пока температура всех частей цепи поддерживается постоянной.* Пока система продолжает оставаться в тепловом равновесии, одна только контактная разность потенциалов термодинамически не может заставить систему быть источником энергии.

В физическом отношении *p-n* переход отличается от обычной батареи тем, что последняя может поддерживать перенос ионов в электролите, что приводит к расходованию одного или обоих электродов в процессе образования новых химических веществ, имеющих более низкую свободную энергию, чем первоначальный материал электродов. В *p-n* переходе этого не происходит, и новые химические вещества не образуются. Если бы доноры и акцепторы могли проходить через *p-n* переход, а кристаллическая решетка играла роль твердого электролита, *тогда* переход обладал бы теми же свойствами, что и обычная батарея, отдавая энергию во внешнюю цепь и уменьшая свою внутреннюю энергию, которая обусловлена начальным разделением доноров и акцепторов в различных областях *p-n* структуры.

Рассмотрим теперь более подробно барьер *p-n* перехода. В частности, мы хотим исследовать распределения избыточных зарядов, которые приводят к распределению потенциала, показанному на рис. 19-2. Диаграмма для плоскостного *p-n* барьера в несколько увеличенном и детализированном виде показана также на рис. 19-3. На этой диаграмме показано, что электроны [1], которые шли справа налево в первые мгновения после того, как образцы *p*-типа и *n*-типа были приведены в соприкосновение, оказываются после установления равновесия занимающими полностью все уровни акцепторных примесей в некоторой области слева от места кон-

---

[1] Необязательно, конечно, эти самые электроны, но эквивалентное число электронов из *p*-области перехода.

такта. Подобным же образом в соответствующей валентной области дырки, двигаясь слева направо, рекомбинируют с электронами в *n*-области перехода и полностью опустошают уровни доноров на некотором протяжении справа от контакта. На рис. 19-3 плюсы в кружочках обозначают опустошенные доноры, а мину-

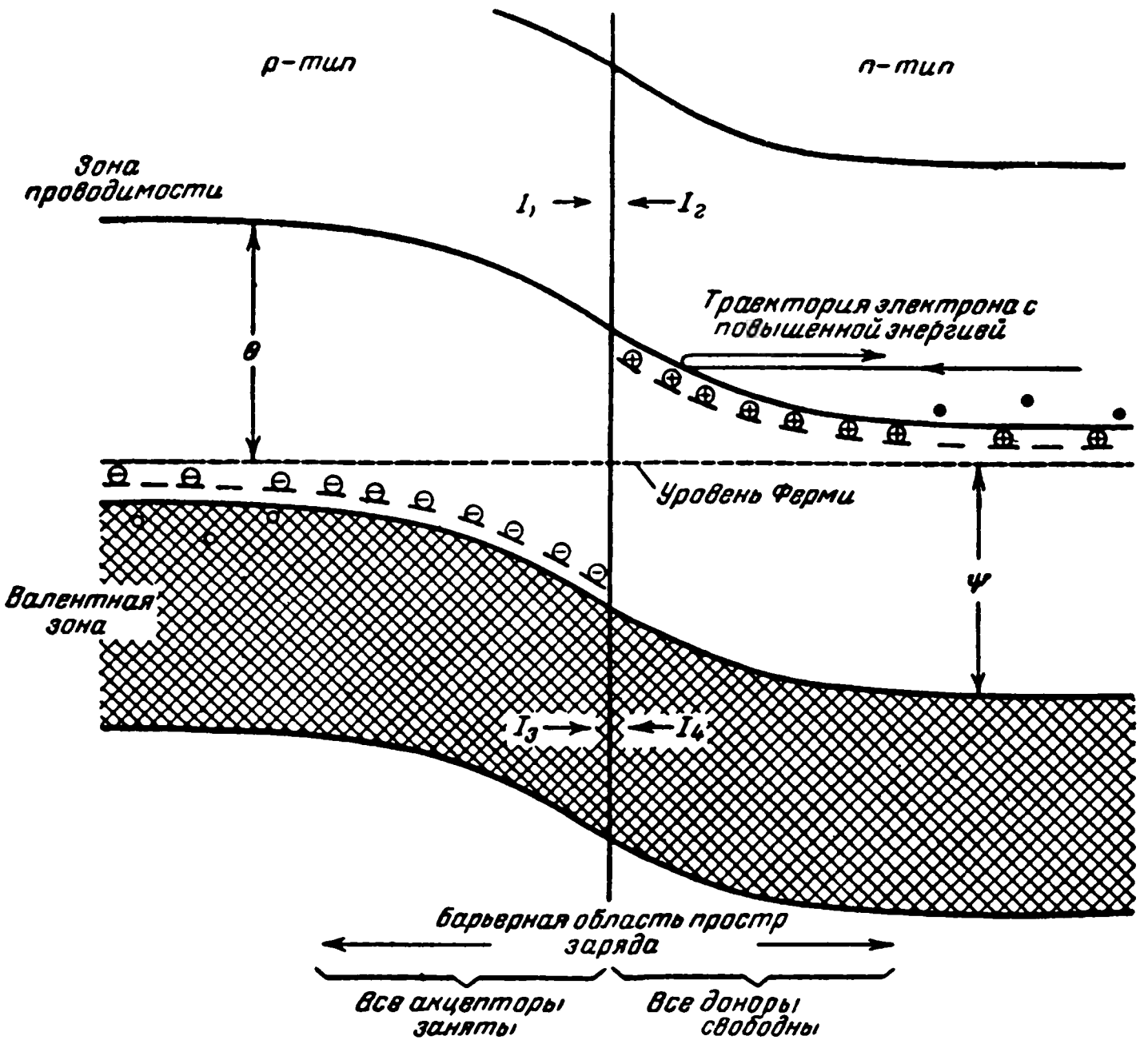


Рис. 19-3 На этой детальной энергетической диаграмме области барьера *p-n* перехода в слоях пространственного заряда показаны ионизированные доноры и акцепторы.

сы в кружочках — заполненные акцепторы. Эти полностью насыщенные ионизованными донорами и акцепторами области, расположенные непосредственно рядом с переходом, образуют двойной слой заряда, который обусловливает контактную разность потенциалов и поле в барьере. Таким образом, барьерная область содержит главным образом фиксированные заряды типа указанных выше свободных доноров или занятых акцепторов и лишь небольшое количество подвижных электронов и дырок, проникающих в барьер на небольшие расстоя-

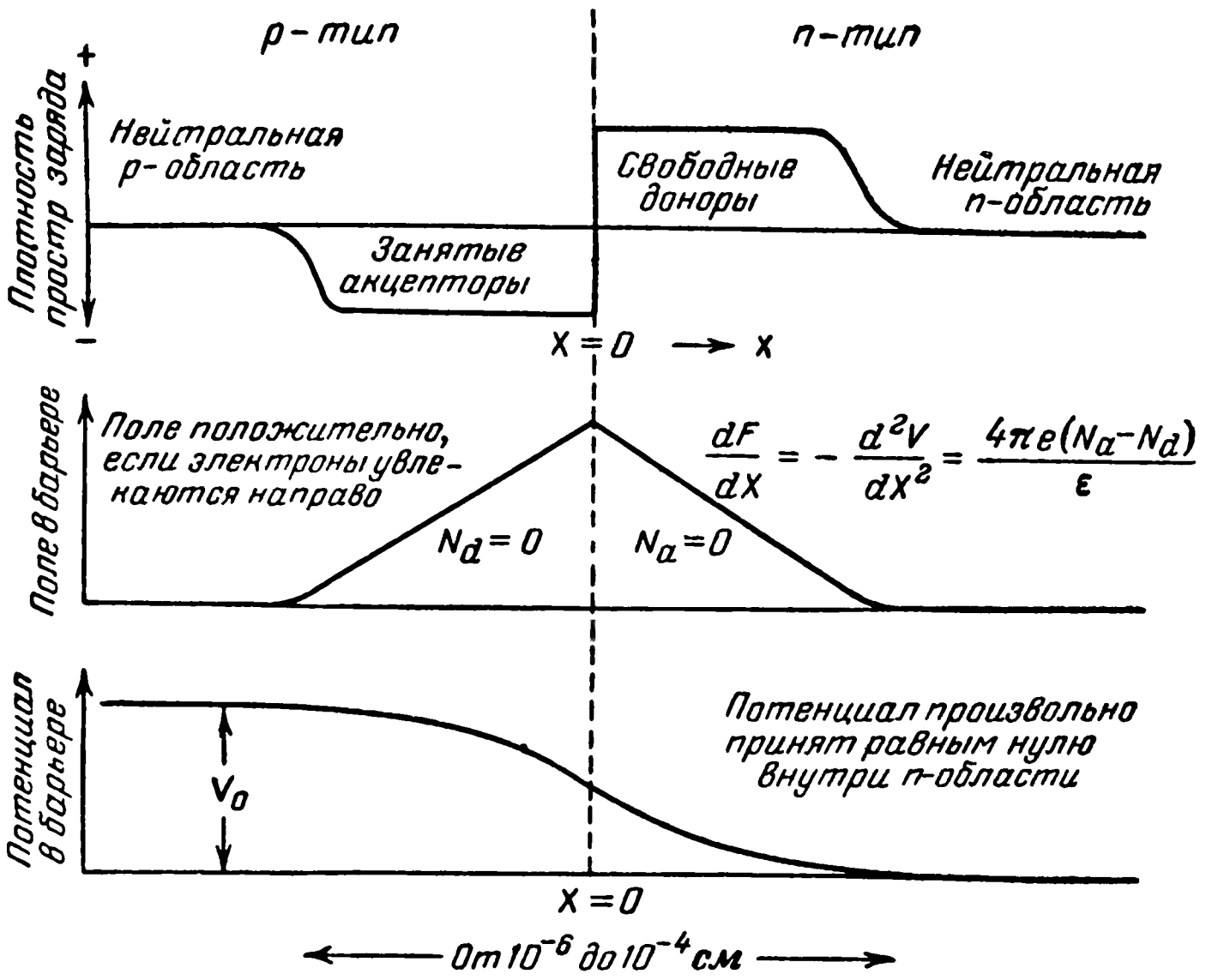


Рис 19-4 Плотность пространственного заряда, поле барьера и распределение потенциала для *p-n* перехода при тепловом равновесии.

ния от его внешних границ[1]. Лишь иногда у случайной частицы — носителя заряда оказывается достаточно энергии, чтобы проникнуть глубоко в барьер, и еще реже у нее хватает энергии на то, чтобы преодолеть поле барьера и пройти сквозь него.

На рис. 19-3 показана траектория электрона проводимости, имеющего энергию, в несколько раз бóльшую, чем средняя тепловая кинетическая энергия, и движущегося вначале справа налево к барьеру. Этот электрон проникает на некоторое расстояние в барьерную область пространственного заряда, останавливается под действием электрического поля барьера и отражается назад внутрь *n*-области. На рис. 19-4 вверху показано распре-

[1] Типичный барьер *p-n* перехода в германии может иметь высоту $^1/_2$ *эв*. Так как тепловая кинетическая энергия основного электрона или дырки при комнатной температуре равна в среднем 3,2 *kT* ($\sim{}^1/_{25}$ *эв*), то нетрудно видеть, что лишь сравнительно небольшое число таких частиц имеет достаточную энергию, чтобы преодолеть хотя бы половину пути в барьере. Доля частиц — носителей заряда, имеющих достаточную энергию, чтобы преодолевать барьер целиком, порядка $e^{-1/2 \,/\, 1/25} = 10^{-5,4}$, или около одной частицы на 250 000.

деление пространственного заряда, образованное ионизованными донорами и акцепторами в барьерной области по обе стороны от перехода, изображенного на рис. 19-3. В середине показано распределение электрического поля, соответствующее этому распределению пространственного заряда. Внизу дано соответствующее распределение потенциала, которое мы уже видели на предыдущих рисунках в виде искривленных краев энергетических зон [1].

Хотя результирующий ток при отсутствии внешнего поля и не идет через переход после того, как достигнуто только что описанное равновесие, не следует думать, что все находится в покое и более ничего не происходит.

Электронные и дырочные токи продолжают идти в обоих направлениях через переход, но всегда величины их равны, а направления противоположны, так что пока частицы продолжают находиться в динамическом тепловом движении, внешне ничего не наблюдается. Эти равновесные токи можно разложить на четыре составляющие, которые мы рассмотрим в отдельности.

1. Ток I_1 неосновных электронов из p-области создается диффузией их через переход слева направо; их энергия соответствует зоне возбуждения (рис. 19-3). Для этих электронов переход играет роль коллектора. По мере того как они уходят, внутри p-области в результате теплового возбуждения постепенно образуются новые неосновные электроны. Используя уравнение непрерывности, можно вычислить величину этой составляющей тока. На рис. 19-5 графически изображена концентрация неосновных электронов в p-области. Слева, на значительном расстоянии от перехода, концентрация

---

[1] При построении графика поля и потенциала может возникнуть затруднение, вследствие того что электростатический потенциал согласно определению представляет собой энергию единичного положительного заряда, тогда как энергетические диаграммы мы привыкли изображать, откладывая на графиках вверх увеличение энергии электронов Этого затруднения можно избежать, определив электростатический потенциал как величину, соответствующую энергии единичного отрицательного заряда. Тогда большим положительным потенциалом был бы такой, при котором энергия электрона увеличилась, а положительное поле имело бы направление, в котором электрон ускоряется этим полем. Мы примем это определение для данного и дальнейших рассуждений. Математически обычно достаточно изменить знаки на обратные, чтобы привести уравнение, содержащее потенциалы, в соотвествие с новым определением.

этих носителей равна n_0, т. е. нормальной величине в условиях теплового равновесия. Однако поскольку переход действует для этих частиц как сток, в пределах нескольких диффузионных длин от перехода появляется спад концентрации неосновных носителей, приводящий

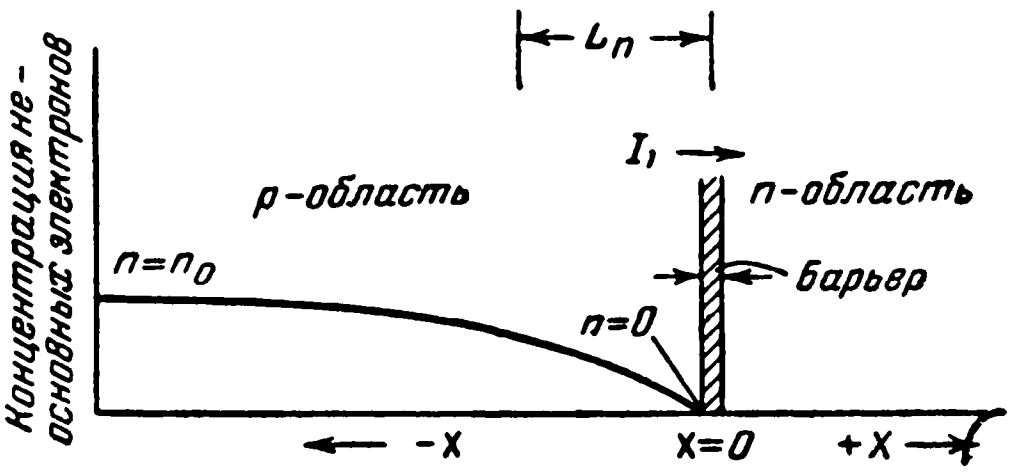


Рис. 19-5 Эта диаграмма концентрации электронов со стороны p-области находящегося в равновесии p-n перехода объясняет возникновение градиента, ответственного за появление составляющей тока I_1, а именно:

$$I_1 = -eD_n \left|\frac{dn}{dx}\right|_{x=0} = -eD_n N_0/L_n.$$

к тому, что n становится равным нулю на самом переходе[1].

В уравнении непрерывности, написанном применительно к этому случаю, $\frac{\partial n}{\partial t}$ отсутствует, так как рассматривается стационарный случай; $\mu_n E_x (\partial n/\partial x)$ исчезает, так как внутри p-области поле отсутствует, а g исчезает, так как никаких других источников или стоков нет, кроме самого перехода, а сток на переходе следует учесть в формулировке граничных условий.

Следовательно,

$$0 = D_n \frac{\partial^2 n}{\partial x^2} - \frac{n - n_0}{\tau_n}, \qquad (19\text{-}1)$$

---

[1] Это не совсем точно. Концентрация этих электронов на переходе равна I_1/ev_n, где v_n — средняя направленная скорость их в месте контакта. Однако эта величина много меньше n_0 и при приближенном рассмотрении ее можно полагать равной нулю. *Прим. ред.*

откуда получается

$$(n - n_0) = Ae^{\pm x/\sqrt{D_n\tau_n}} \; см^{-3},$$

где только положительная экспонента дает верное решение для отрицательных значений x. При $x=0$ $n=0$; отсюда

$$A = -n_0 \text{ и } n = n_0\left(1 - e^{x/\sqrt{D_n\tau_n}}\right) \; см^{-3}.$$

Градиент dn/dx на переходе равен:

$$\left.\frac{dn}{dx}\right|_{x=0} = -\frac{n_0}{\sqrt{D_n\tau_n}} \frac{см^{-3}}{см}. \qquad (19\text{-}2)$$

Число электронов, диффундирующих в секунду через переход слева направо, в $-D_n$ раз больше численного значения этого градиента, а плотность диффузионного тока в $-e$ раз больше полученной в результате величины. Следовательно,

$$I_1 = -\frac{eD_n n_0}{\sqrt{D_n\tau_n}}.$$

Используя статистику Ферми для той части p-области, которая удалена от перехода, получаем:

$$n_0 = 2UT^{3/2}e^{-e\theta/kT} \; см^{-3},$$

где величина $\theta (= V_g - V_{Fp})$ показана на рис. 19-2 в предположении, что θ в несколько раз больше, чем kT, так что n_0 определяется экспоненциальным участком („хвостом“) функции Ферми.

Отсюда

$$I_1 = -\frac{eD_n}{\sqrt{D_n\tau_n}} 2UT^{3/2}e^{-e\theta/kT} \; а/см^2. \qquad (19\text{-}3)$$

2. Ток I_2 основных электронов идет справа налево через барьер; эти электроны имеют энергию, соответствующую зоне возбуждения. Ток I_2 создается сравнительно небольшим числом возбужденных электронов n-области полупроводника, имеющих достаточную энергию для преодоления противодействующего поля внутри барьера. Из этой группы электронов лишь некоторая

часть в каждый момент времени движется по направлению к барьеру, и из этой части лишь немногие электроны могут проникнуть в барьер прежде, чем они испытают новое столкновение.

Некоторые из этих электронов теряют энергию в результате столкновений до прохода сквозь барьер и возвращаются в n-область структуры. Другие же, испытав столкновения внутри барьера даже без потери энергии, тем не менее отражаются обратно и движутся в том направлении, откуда они пришли. Однако их места занимают другие электроны, которые, преодолев часть пути через барьер, приобретают благодаря столкновениям достаточную энергию, чтобы пройти барьер до конца.

Таким образом, общая картина составляющей тока I_2 очень сложна и проще всего аналитически установить вид этой составляющей, приняв, что в равновесии I_2 должно быть равно и противоположно по знаку I_1.

Следовательно,

$$I_2 = \frac{eD_n}{\sqrt{D_n\tau_n}} 2UT^{3/2}e^{-e\theta/kT} \quad a/см^2. \qquad (19\text{-}4)$$

Когда эти электроны появляются в p-области полупроводника, они обладают в основном тепловыми скоростями и становятся, таким образом, *неосновными* носителями заряда в области слева от барьера.

3. Ток I_3 положительных дырок идет сквозь барьер слева направо; энергии этих дырок соответствуют валентной зоне. Имеются в виду, очевидно, положительные дырки, занимающие достаточно низкие энергетические уровни в p-области, так что барьер не является препятствием для их движения из p- в n-область. (Дырка имеет *более высокую* энергию, если она занимает *более низкий* уровень электронной энергии, точно так же как пузырек в жидкости имеет тем бóльшую энергию, чем глубже он находится в жидкости. Чтобы перенести электрон на более высокий энергетический уровень, необходимо совершить работу; точно так же необходимо совершить работу, чтобы сдвинуть дырку вниз на энергетической диаграмме. Именно поэтому дырок меньше на более низких энергетических уровнях, что аналитически выражается в статистике Ферми.) Эта составляющая тока при тепловом равновесии должна

быть равна и противоположна по знаку составляющей I_4, которую мы рассмотрим в следующем пункте. Уравнение для I_3:

$$I_3 = \frac{eD_p}{\sqrt{D_p\tau_p}} 2UT^{3/2} e^{-e\Psi/kT} \ a/см^2, \qquad (19\text{-}5)$$

где величина Ψ (равная V_{Fn}) показана на рис. 19-3.

4. Ток I_4, созданный положительными дырками, образующимися в n-области и представляющими собой здесь неосновные носители зарядов, идет вследствие диффузии справа налево через барьер и под барьером в валентной зоне. Эта составляющая тока может быть выведена точно таким же методом, как и составляющая I_1.

Следовательно,

$$I_4 = -\frac{eD_p}{\sqrt{D_p\tau_p}} 2UT^{3/2} e^{-e\Psi/kT} \ a/см^2. \qquad (19\text{-}6)$$

При тепловом равновесии токи I_1 и I_2 равны по величине и противоположно направлены, взаимно уничтожая друг друга. То же можно сказать о токах I_3 и I_4. Суммарный ток на переходе поэтому равен нулю, если не приложено внешнее напряжение, хотя в переходе и вблизи него имеются сильное собственное электрическое поле и высокий градиент концентрации носителей.

19-2. ОБРАТНЫЙ И ПРЯМОЙ ТОКИ ЧЕРЕЗ *p-n* ПЕРЕХОДЫ

Рассмотрим теперь эти четыре составляющие тока и зависимость их от обратных и прямых смещений напряжения, приложенных к переходу. На рис. 19-6 показана диаграмма энергетических уровней для p-n перехода, находящегося под действием обратного смещающего напряжения V_a, приложенного в таком направлении, чтобы сделать n-область положительной, а p-область — отрицательной относительно друг друга. Направление этого напряжения таково, что соответствующая барьеру разность потенциалов увеличивается сравнительно с контактной. Наличие этого смещения приводит к неравновесному состоянию системы, причем уровень Ферми уже изображается на диаграмме не горизонтальной прямой, а смещен в вертикальном направлении между n- и p-областями структуры на величину V_a.

Поскольку область пространственного заряда барьера оказывает большое сопротивление обратному току, практически все приложенное напряжение падает на барьере. Сильное электрическое поле внутри барьера выталкивает заряженные частицы из области пространственного заряда, так что местная концентрация подвижных носителей внутри барьера оказывается малой

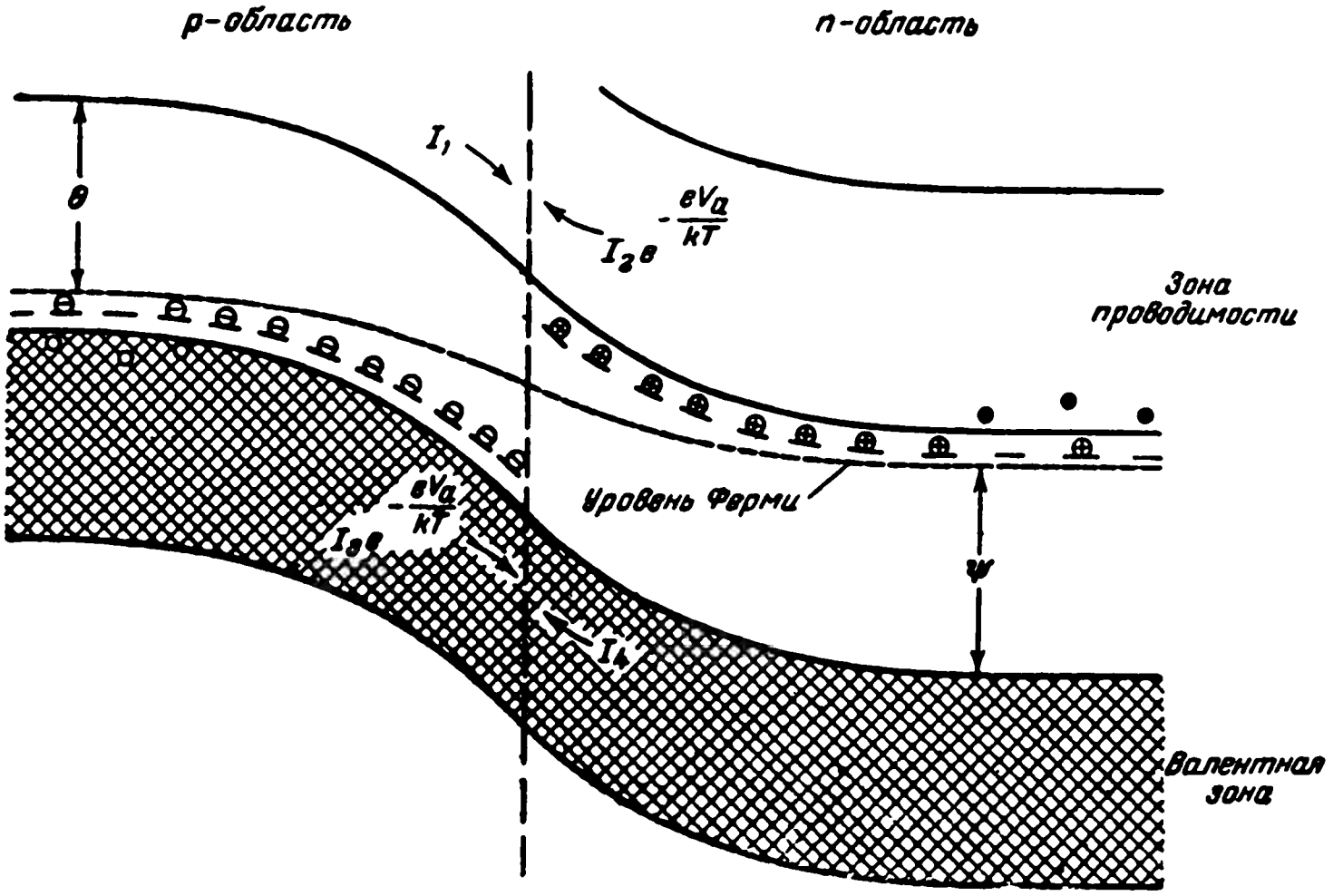


Рис 19-6. Энергетическая диаграмма p-n перехода при обратном смещении $\sim 0{,}5$ *в*. Показаны непостоянство энергии уровня Ферми, увеличение высоты потенциального барьера (на величину V_a) и уменьшение составляющих тока I_2 и I_3.

и местное удельное сопротивление внутри барьера может оказаться на несколько порядков выше, чем собственное сопротивление полупроводника, из которого изготовлен прибор. Причина того, почему при обратном смещении скачок потенциала на барьере и поле в барьере возрастают, заключается в «обнажении» большего числа ионизованных доноров в n-области и большего числа заполненных (ионизованных) акцепторов в p-области перехода. Вследствие этого области пространственного заряда расширяются при приложении обратного смещения. Расстояние между «обкладками» эквивалентного конденсатора, таким образом, возрастает, а емкость барьера уменьшается согласно уравнению (6-2) Шоттки.

Из графика ясно видно, что увеличение высоты барьера, обусловленное обратным смещением, уменьшает составляющие тока I_2 и I_3. Например, электрон в зоне проводимости n-области, стремясь преодолеть барьер в направлении справа налево, должен в этом случае иметь энергию $e(\theta + V_a)$ над уровнем Ферми в n-области. Это обстоятельство приводит к сокращению количества электронов, имеющих энергию, достаточную для того, чтобы участвовать в создании составляющей тока I_2.

Если заменить θ в экспоненте уравнения (19-4) на $(\theta + V_a)$, чтобы уравнение (19-4) соответствовало условиям при смещении, то, очевидно, составляющая I_2 уменьшится в $e^{-eV_a/kT}$ раз от ее равновесного значения. Если V_a более чем в 4 раза превышает kT ($\sim 0{,}1$ *эв*), то I_2 уменьшится до величины, меньшей 2% ее равновесного значения. То же и по той же причине произойдет с I_3.

Однако такое уменьшение величин I_2 и I_3 не меняет составляющих диффузионного тока I_1 и I_4, значения которых продолжают соответствовать уравнениям (19-3) и (19-6). Поэтому если I_2 и I_3 становятся при приложенном достаточно большом обратном смещении пренебрежимо малыми, то сумма I_1 и I_4 будет представлять собой *обратный ток насыщения* этого перехода Таким образом,

$$I_{\text{нас}} = -2eUT^{3/2}\left(\frac{D_n}{\sqrt{D_n\tau_n}}e^{-e\theta/kT} + \frac{D_p}{\sqrt{D_p\tau_p}}e^{-e\psi/kT}\right) \quad a/см^2. \tag{19-7}$$

Уравнение (19-7) дает представление о природе и составе обратного тока насыщения, но его форма не позволяет проводить простые вычисления. Используя следующие замены:

$$\left.\begin{aligned} n &= 2UT^{3/2}e^{-e\theta/kT}; \\ p &= 2UT^{3/2}e^{-e\psi/kT} \end{aligned}\right\} \text{— концентрации частиц;}$$

$D = \frac{kT\mu}{e}$ — соотношение Эйнштейна;

$pn = n_i^2$ — из уравнения (17-28);

$\sigma_i = en_i(\mu_n + \mu_p)$ — из уравнения (17-29);

$b = \frac{\mu_n}{\mu_p}$ — отношение подвижности электронов к подвижности дырок;

$\sqrt{D\tau} = L$ — определение диффузионной длины,

можно показать, что (19-7) эквивалентно

$$I_{\text{нас}} = -\frac{kT}{e} \cdot \frac{b}{(1+b)^2} \sigma_i^2 \left(\frac{1}{\sigma_n L_p} + \frac{1}{\sigma_p L_n} \right) \text{ а/см}^2. \quad (19\text{-}8)$$

Напомним еще раз, что этот обратный ток насыщения создается потоком частиц, которые представляют собой *неосновные носители* в тех областях, из которых они приходят. Пройдя переход, эти частицы становятся основными носителями и теряются в большой массе основных носителей области, в которую они пришли. Обратный ток насыщения есть не что иное, как диффузионный ток, являющийся результатом наличия градиентов концентрации неосновных носителей в переходе и вблизи него и, следовательно, не зависящий от приложенного напряжения. Этим объясняется форма кривой обратного тока, показанная на рис. 5-15 для плоскостного *p-n* диода. Область насыщения начинается при напряжении достаточно большом ($\sim$0,1 *в*), чтобы уменьшить составляющие тока I_2 и I_3 до пренебрежимо малых величин, и продолжается до напряжения настолько высокого ($\sim$150 *в* в этом частном случае), что в результате наступает явление пробоя перехода.

Во многих случаях, когда насыщения на обратной характеристике не наблюдается, объяснения с помощью приведенных выше теоретических соображений дать не представляется возможным. Иногда ненасыщение характеристики обратного тока выпрямителя приписывается краевой утечке через переход по поверхности полупроводника. Эти явления особенно заметны, когда полупроводник недостаточно хорошо протравлен и на поверхности после механической его обработки остались дефекты. Во влажной атмосфере незащищенные элементы *p-n* перехода германия могут иметь значительную утечку, как бы тщательно они ни были приготовлены. В этом случае можно предположить, что причиной утечки являются химические и электролитические процессы в адсорбированном на поверхности слое, имеющем толщину в несколько молекул. В выпрямителях,

использующих контакт металла с полупроводником, отсутствие насыщения на характеристике обратного тока является результатом уменьшения величины работы выхода металла сильным полем у его поверхности, созданным приложенным обратным напряжением. При этом по мере увеличения обратного напряжения ток на-

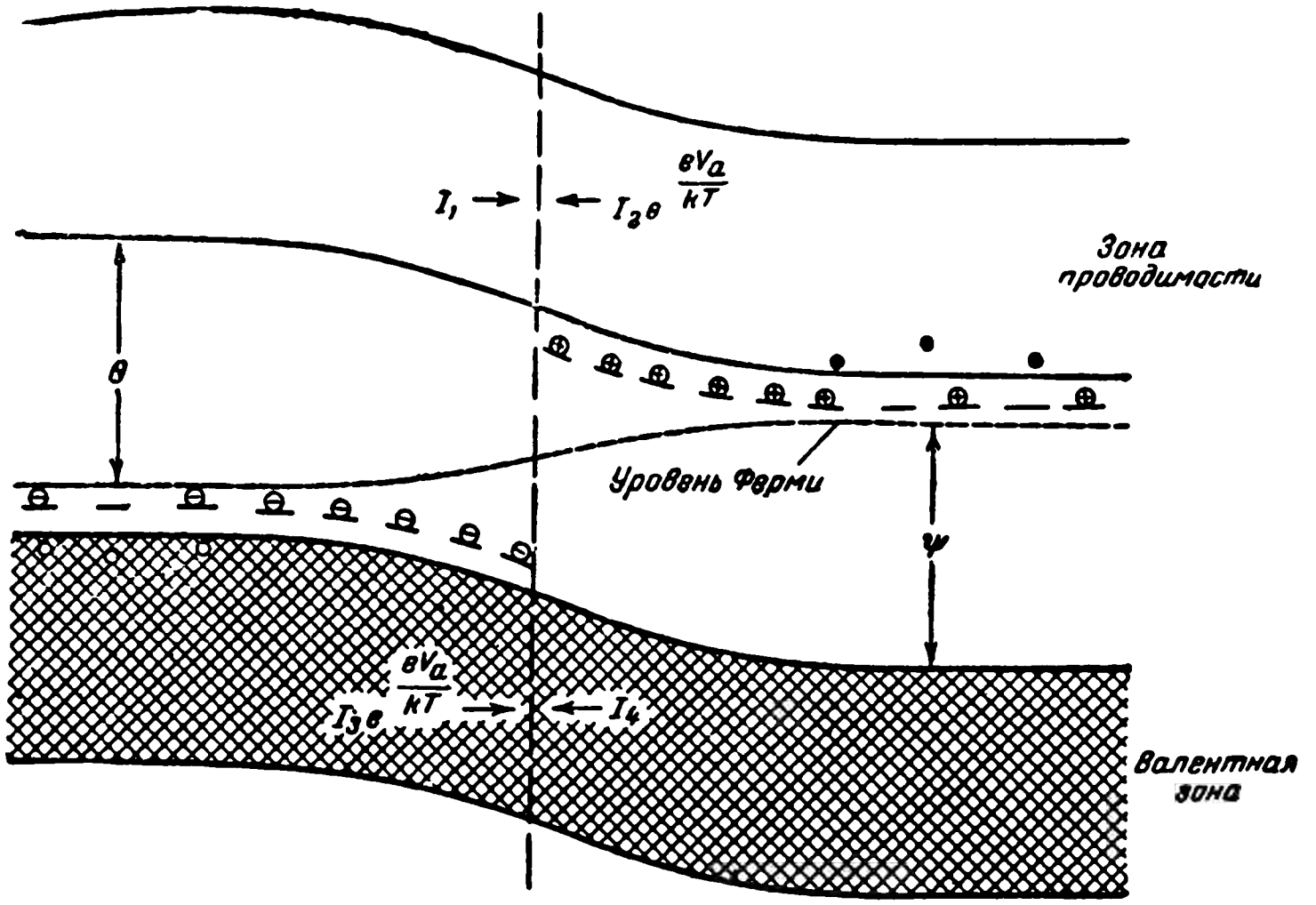


Рис. 19-7. Энергетическая диаграмма p-n перехода при прямом смещении $\sim 0{,}25$ $в$ Показаны непостоянство энергии уровня Ферми, уменьшение высоты потенциального барьера (на величину V_a) и соответствующее увеличение составляющих тока I_2 и I_3 на множитель $e^{eV_a/kT}$.

сыщения может увеличиваться и вследствие возникновения тока автоэлектронной эмиссии, возрастающего по мере увеличения V_a. Это явление рассмотрено в следующей главе.

Перейдем теперь к рассмотрению прямого тока, проходящего через барьер. Рис. 19-7 соответствует переходу, находящемуся под прямым смещением (n-область отрицательна относительно p-области) вследствие приложенного напряжения, близкого к $^1/_2V_g$. При таком направлении приложенного напряжения высота равновесного барьера уменьшается на величину V_a.

Как и прежде, равновесие системы нарушается и уровень Ферми в пределах барьера подвергается верти-

кальному смещению на величину V_a. Направление этого смещения таково, что контактная разность потенциалов частично уменьшается и высота барьера понижается. Очевидно, что число основных носителей по обе стороны перехода, обладающих достаточной энергией для того, чтобы преодолеть бырьер, увеличивается, так как в этом случае для преодоления барьера требуется меньшая энергия. Составляющие тока I_2 и I_3 в этом случае увеличиваются в $e^{eV_a/kT}$ раз, тогда как составляющие I_1 и I_4 остаются без изменения. Отсюда плотность общего тока через переход

$$I_{\text{общ}} = (I_1 + I_4) + (I_2 + I_3)\, e^{eV_a/kT}$$

или, так как

$$I_2 + I_3 = -(I_1 + I_4)$$

и

$$I_1 + I_4 = I_{\text{нас}},$$

то

$$I_{\text{общ}} = I_{\text{нас}}\,(1 - e^{eV_a/kT})\ \ a/см^2, \qquad (19\text{-}9)$$

где $I_{\text{нас}}$ дано в уравнении (19-8), а V_a считается положительной величиной для прямого смещения. (Поскольку $I_{\text{нас}}$ само по себе отрицательно, прямое смещение дает ток положительного знака.)

Таким образом, видно, что прямой ток p-n перехода создается потоками частиц, являющихся основными носителями в областях, из которых они поступают, и становящихся неосновными носителями после того, как они пересекли барьер. Причиной того, что они идут сквозь барьер, является, как и раньше, не электрическое поле, а диффузия в направлении градиентов концентрации от одной стороны барьера к другой. Приложенное напряжение лишь изменяет форму зависимости этих градиентов от расстояния, делая ее более благоприятной для диффузионного тока, идущего в прямом направлении.

Уравнение (19-9) известно как **уравнение выпрямления.** Оно определяет форму кривой зависимости тока от напряжения в вентиле с p-n переходом как при прямом, так и при обратном смещении. Рис. 19-8 представ-

ляет собой графическое изображение этой зависимости для германиевого p-n перехода, имеющего $I_{нас}=$ $=200$ *мка/см*^{2} и площадь 1 *мм*^{2}. Поразительно сходство этой характеристики с измеренной зависимостью для

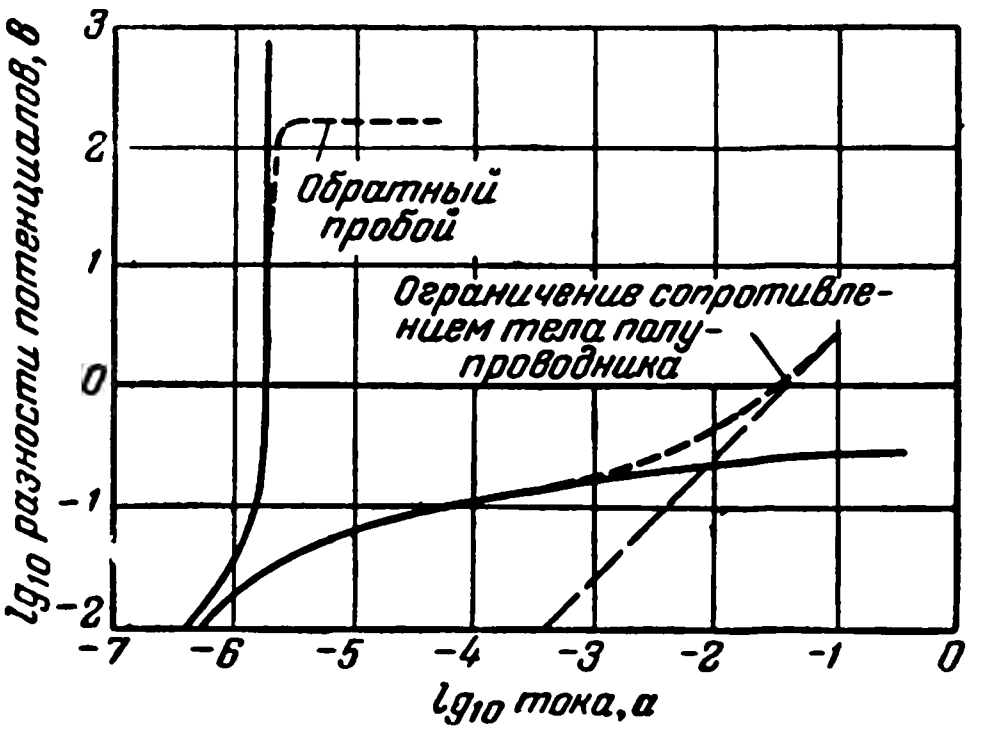


Рис 19-8 Характеристика идеального p-n перехода, рассчитанная из уравнения выпрямителя (сплошная линия), совпадает с экспериментальной характеристикой выпрямителей с p-n переходами

реального вентиля с p-n переходом, показанной на рис. 5-15; исключением являются области с высоким обратным напряжением, где имеет место пробой, и с высоким прямым током, где появляется ограничивающий фактор в виде сопротивления [Л. 1] тела самого выпрямителя.

19-3. ИСПОЛЬЗОВАНИЕ p-n ПЕРЕХОДОВ В КАЧЕСТВЕ ЭМИТТЕРОВ

До сих пор, приводя для иллюстрации примеры p-n переходов, мы имели в виду переходы с примерно равным количеством донорных и акцепторных примесей по обе стороны от контакта, так что области пространственного заряда в барьере и кривая потенциала симметричны по обе стороны от перехода. В результате величины θ и ψ на рис. 19-3 приблизительно равны и, если не учитывать различия в величинах подвижностей электронов и дырок, составляющие тока I_2 и I_3 оказываются примерно равными. Это как раз те составляю-

щие тока, которые при прямом смещении V_a, соответственно умноженные на множитель $e^{-eVa/kT}$ практически составляют весь прямой ток перехода, если смещение превышает 0,1 *в*.

На практике редко встречаются случаи с одинаковым количеством примесей по обе стороны *p-n* перехода. В действительности при изготовлении эмитерного перехода заранее предопределяется неравное количество примесей. Так как работа плоскостного транзистора зависит от введения и переноса через базовый слой частиц, которые являются *неосновными носителями заряда* в базовой области, необходимо при разработке эффективного транзистора учитывать, что эмитерный ток должен быть в основном током неосновных носителей, эмитируемых в базу, а не током основных носителей, исходящих из базы. Это требование будет удовлетворено, если в эмитерную область эмитерного *p-n* перехода внести большее количество примесей, чем в базовый слой.

На рис. 19-9 показан *p-n* переход (например, эмитерный переход в *p-n-p* транзисторе), в котором *p*-область имеет больше акцепторов в 1 $см^3$, чем *n*-область имеет доноров в 1 $см^3$. Неодинаковое количество примесей приводит к асимметрии барьера, при которой бóльшая часть падения потенциала на переходе в основном происходит с той стороны, где концентрация примесей меньше, как показано на рисунке. Область пространственного заряда с *p*-стороны у́же, чем с *n*-стороны, так как бóльшая концентрация примесей дает и более плотный пространственный заряд ионизованных примесей. Далее, расстояние уровня Ферми от границы зоны возбуждения с правой стороны от барьера значительно больше расстояния, на котором он находится над границей валентной зоны с левой стороны от барьера; так как состояние области с меньшей концентрацией примесей ближе к состоянию беспримесного полупроводника, уровень Ферми лежит ближе к середине энергетической щели в *n*-области. Поэтому величина θ больше, чем ψ, и составляющая I_3 в выражении для прямого тока больше, чем составляющая I^*_2. Прямой ток перехода

---

* Это утверждение предполагает, что D_n и τ_n не являются совершенно несравнимыми с D_p и τ_p. Для германиевых и кремниевых транзисторов это предположение справедливо.

поэтому в основном представляет собой ток положительных дырок, идущих в n-область под барьером в валентной зоне. В этой области дырки становятся неосновными носителями и их дальнейшая диффузия направо и то, как они собираются близрасположенным

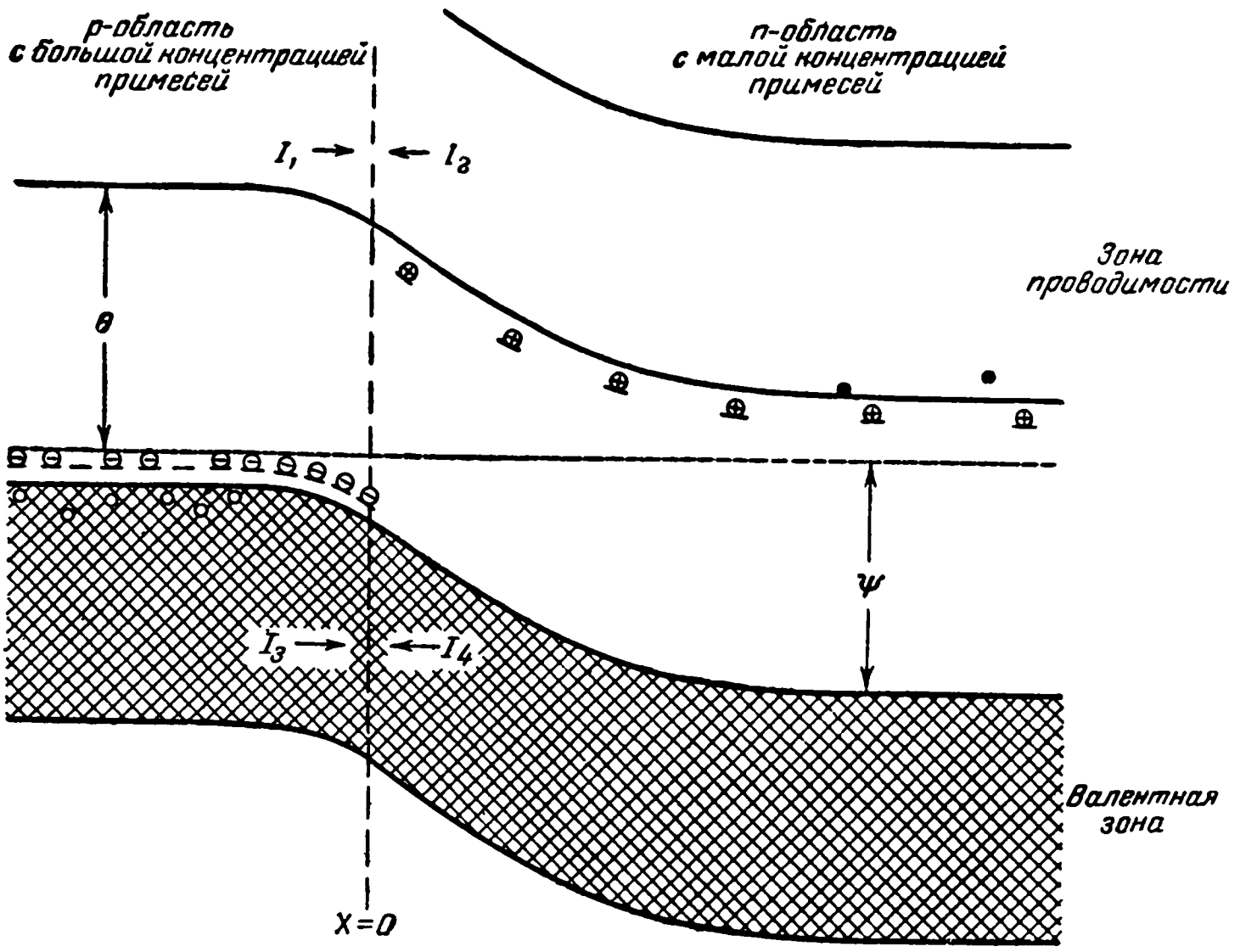


Рис 19-9. На энергетической диаграмме p-n перехода с несимметричным активированием уровень Ферми расположен ближе к середине энергетической щели в n-области, а $\theta > \Psi$. Соответственно $I_3 > I_2$ и при прямом смещении ток перехода в основном представляет собой дырочный ток (эмиссию неосновных носителей в область n-типа).

коллекторным переходом, составляют существо действия транзистора.

Количественно, если θ больше, чем Ψ, примерно на 5 kT (приблизительно на 0,12 $в$), то прямой ток составит примерно 99% положительного дырочного тока. Используя статистику Ферми, можно рассчитать, где пройдет уровень Ферми относительно границ зон в p- и n-областях с разных сторон от барьера при любых концентрациях примесей N_a и N_d, и таким образом определить θ (равные $V_g - V_{Fp}$ в p-области) и Ψ (равное V_{Fn} в n-области). И, наоборот, если известно, какие значе-

ния θ и ψ необходимы для получения требуемого отношения неосновного тока к основному, можно вычислить соответствующие пары значений N_a и N_d.

При обсуждении эмиттеров и их свойств часто используется выражение для *эффективности эмиссии неосновных носителей*, обозначаемое γ. Имеется в виду отношение тока неосновных носителей к общему прямому току, проходящему через эмитерный переход. Когда этот переход находится под достаточным прямым смещением, так что составляющие I_1 и I_4 пренебрежимо малы по сравнению с составляющими I_2 и I_3 вследствие возрастания последних за счет множителя прямого смещения $e^{eV_a/kT}$, то по определению

$$\gamma = \frac{I_3}{I_3 + I_2} = \frac{1}{1 + \frac{\sigma_n L_p}{\sigma_p L_n}}, \qquad (19\text{-}10)$$

где принято, что p-область эмитирует неосновные носители в n-область.

Можно получить дополнительные сведения о поведении потоков частиц в прямом токе внутри и вблизи p-n перехода с прямым смещением, рассмотрев концентрации основных и неосновных носителей на различных расстояниях от этого перехода. Например, рассмотрим переход при прямом смещении, подобный тому, который показан на рис. 19-9. Он имеет настолько высокую концентрацию примесей в p-области, что составляющей I_2 в выражении для прямого тока можно пренебречь. Поэтому можно рассматривать только положительный дырочный ток I_3, идущий слева направо в валентной зоне. Для этого случая кривые концентраций различных основных и неосновных носителей по обе стороны барьера показаны на рис. 19-10. Обратите внимание, что часть этой задачи уже была рассмотрена в примере 3 § 18-3, причем нижняя правая часть рис. 19-10 точно такая же, как и весь рис. 18-2, за исключением логарифмической шкалы ординат.

Необходимо особо указать на то, что справа от барьера ток переносится почти целиком как диффузионный ток положительных дырок, которые диффундируют через барьер в валентной зоне и после этого диффундируют направо в направлении градиента концентрации (положительных) дырок.

По мере того как эти частицы удаляются от барьера, их концентрация уменьшается вследствие рекомбинации, а градиент концентрации становится меньше. Таким образом, по мере удаления направо от барьера все бóльшая часть тока цепи создается основными носите-

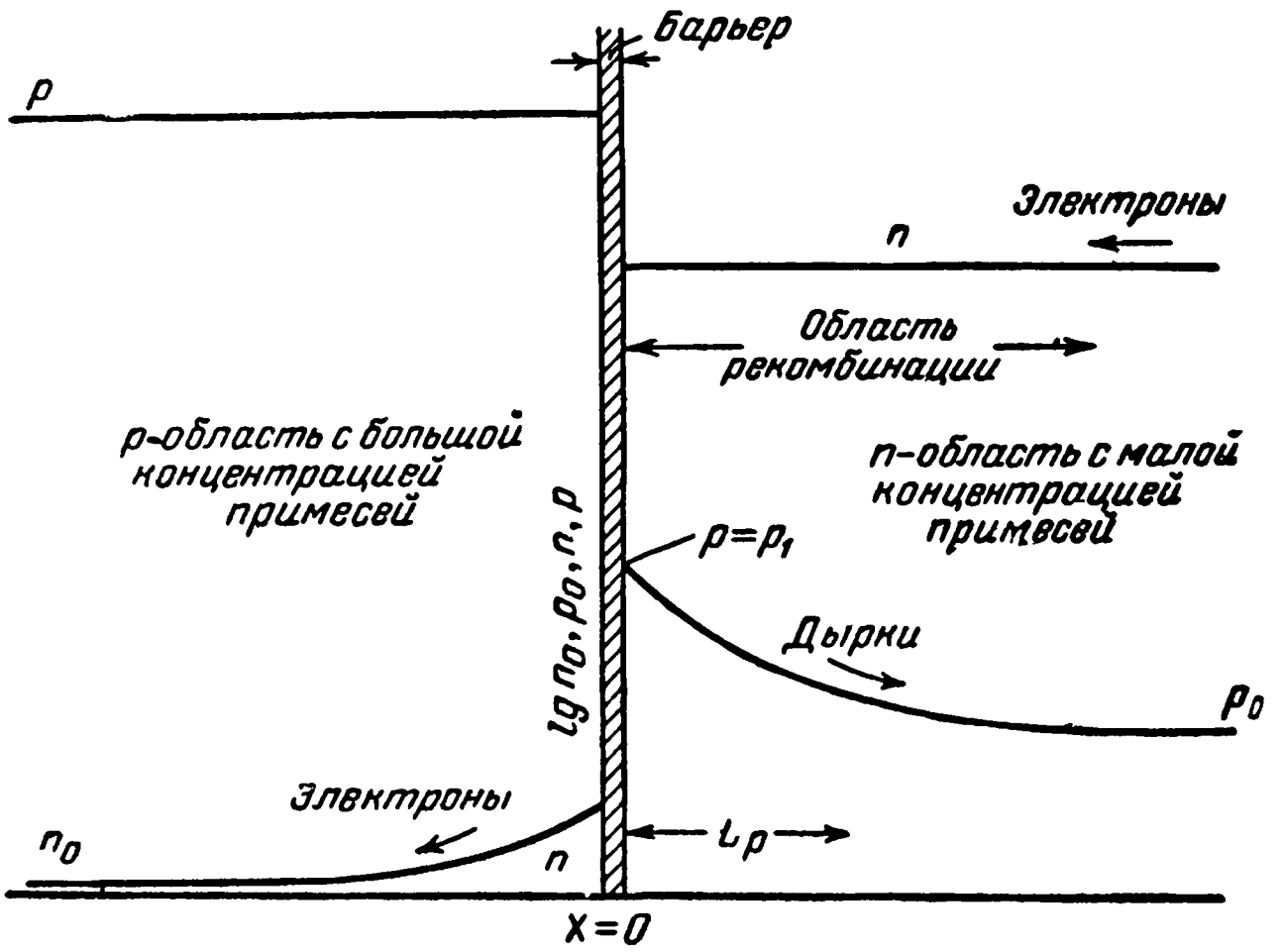


Рис. 19-10. Если *p-n* переход находится под прямым смещением, прямой ток переносится частицами, которые становятся неосновными носителями, после того как они прошли через переход. Вблизи перехода эти частицы создают ток вследствие диффузии в направлении градиентов концентрации неосновных носителей. На некотором расстоянии от перехода ток создается основными носителями, увлекаемыми полем в полупроводнике из-за внешнего смещения

лями — электронами, движущимися справа налево в зоне проводимости. Эти основные носители перемещаются по направлению к барьеру, чтобы занять места уже рекомбинировавших. На достаточно большом расстоянии от барьера вправо, скажем на расстоянии 3 или 4 диффузионных длин положительных дырок, где p спадает практически до p_0, концентрация неосновных носителей становится постоянной, а диффузионный ток положительных дырок делается равным нулю и почти весь ток цепи правее этого места создается основными частицами — электронами. Эти электроны создают здесь *ток проводимости*, обусловленный малым падением напря-

жения IR в правой части полупроводника. Этот электронный ток проводимости гораздо больше, чем дырочный ток проводимости, так как в n-области гораздо больше электронов, чем дырок. Поэтому, используя в гл. 18 уравнение непрерывности для подобного случая, мы были вправе предположить, что в n-области практически нет поля, поскольку это касалось положительных дырок. Действительно, хотя в принципе *существует* небольшой ток проводимости положительных дырок справа от области рекомбинации, показанной на рис. 19-10, этот ток пренебрежимо мал сравнительно с диффузионным током положительных дырок, идущим внутри рекомбинационной области.

Если бы переход был более симметричным, конечно, пришлось бы учитывать также электронный диффузионный ток слева от барьера. Однако в рассматриваемом случае, в силу того что в p-области велика концентрация примесей, n_0, n и dn/dx все слишком малы, чтобы в значительной степени увеличивать ток с левой стороны от барьера, и практически весь ток переносится в этой области с помощью основных дырок, движущихся в электрическом поле с градиентом IR в p-области.

19-4. ЕМКОСТЬ БАРЬЕРА

Мы уже несколько раз отмечали, что поведение p-n перехода подобно поведению конденсатора, имеющего в качестве обкладок полупроводники n- и p-типа, разделенные тонким слоем диэлектрика с очень высоким сопротивлением — самим барьером. Существенное отличие емкости перехода от обычного конденсатора заключается в том, что в конденсаторе заряды на противоположных поверхностях металлических пластин располагаются чрезвычайной тонким слоем, тогда как в p-n переходе эти слои имеют определенную толщину и заряд распределен по всему объему областей пространственного заряда по обе стороны от контакта. Емкость перехода на единицу площади перехода можно вычислить, если известны концентрация доноров и акцепторов в областях пространственного заряда n- и p-типа и диэлектрическая проницаемость полупроводника.

Используя уравнение Пуассона, можно рассчитать суммарную ширину областей пространственного заряда при заданных концентрациях доноров и акцепторов, не-

обходимую для получения разности потенциалов, существующей на барьере. В результате получим:

$$\textit{Толщина барьера} = 1{,}05 \cdot 10^{-6} \left(\frac{\varepsilon (U_0 - U_a)}{2\pi e (N_a + N_d)} \right)^{1/2} \times$$

$$\times \left\{ \left(\frac{N_a}{N_d} \right)^{1/2} + \left(\frac{N_d}{N_a} \right)^{1/2} \right\} \textit{см}, \qquad (19\text{-}11)$$

где ε — диэлектрическая проницаемость полупроводника (16 для германия); U_0 — высота потенциального барьера у контакта при отсутствии смещения и U_a — величина приложенного внешнего напряжения смещения в случае его наличия, причем U_a положительно, если смещение прямое, и отрицательно, если смещение обратное.

Цифровой множитель в уравнении (19-11) соответствует измерению U_0 и U_a в вольтах, а e — в кулонах. В частном случае, если $N_d = N_a = N$, барьер симметричен и

$$\textit{Толщина симметричного барьера} =$$

$$= 1{,}05 \cdot 10^{-6} \left(\frac{\varepsilon (U_0 - U_a)}{\pi e N} \right)^{1/2} \textit{см}. \qquad (19\text{-}11\text{а})$$

С другой стороны, если барьер сильно не симметричен при, скажем, $N_a \gg N_d$, как показано на рис. 19-9, то уравнение (19-10) принимает форму

$$\textit{Толщина барьера при } N_a \gg N_d =$$

$$= 1{,}05 \cdot 10^{-6} \left(\frac{\varepsilon (U_0 - U_a)}{2\pi e N_d} \right)^{1/2} \textit{см}. \qquad (19\text{-}11\text{б})$$

Из уравнений (19-11), (19-11а) и (19-11б) следует, что толщина барьера зависит от смещения, как говорилось ранее, а также от концентрации примесей по обе стороны перехода.

Емкость барьера на 1 *см* его площади равна:

$$C = \frac{\varepsilon}{4\pi \times \textit{ширина барьера}} \textit{см}.$$

Для симметричного барьера по уравнению (19-11а)

$$C = 1{,}05 \left(\frac{\varepsilon e N}{16\pi (U_0 - U_a)} \right)^{1/2} \textit{мкф/см}^2. \qquad (19\text{-}12\text{а})$$

Для асимметричного барьера по уравнению (19-11б)

$$C = 1{,}05 \left(\frac{\varepsilon e N_d}{8\pi (U_0 - U_a)} \right)^{1/2} \text{мкф/см}^2. \qquad (19\text{-}12б)$$

Уравнение (19-12б) имеет тот же вид, что и уравнение Шоттки (6-2) для емкости вентиля с контактом металл — полупроводник. В отношении емкости переход между двумя полупроводниками, из которых один имеет очень высокую концентрацию примеси, эквивалентен выпрямляющему контакту металла с полупроводником.

Все сказанное выше основано на модели p-n перехода, в котором изменение содержания примесей от концентрации акцепторов N_a к концентрации доноров N_d происходит на расстоянии одного-двух атомных диаметров (ступенчатый переход, резкий переход). Такой переход от p-области к n-области не осуществляется при обычных процессах изготовления p-n переходов путем выращивания или диффузии; в этих случаях он растягивается на многие сотни и даже тысячи атомных диаметров.

Если расстояние, на котором происходит изменение активировки от полностью акцепторной к полностью донорной мало по сравнению с толщиной области пространственного заряда барьера, приведенный выше анализ останется верным. Однако если этот переход настолько расширяется, что он занимает пространство, сравнимое или даже большее, чем толщина области пространственного заряда барьера, то переход принято называть «постепенным» или «плавным» и в выражении для емкостей приходится вносить изменения.

Одним из типов такого плавного перехода, который легко поддается расчету и который довольно хорошо воспроизводит встречающиеся в практике случаи, является переход, где величина $(N_a - N_d)$ изменяется линейно.

Для такого плавного перехода:

$$C = \frac{1{,}05}{4} \left(\frac{\varepsilon^2 e G}{3\pi^2 (U_0 - U_a)} \right)^{1/3} \text{мкф/см}^2, \qquad (19\text{-}13)$$

где G — градиент разности $(N_a - N_d)$ в области перехода.

Это выражение отличается от выражений для резких переходов тем, что здесь величина емкости изменяется обратно пропорционально кубическому корню из величины смещения.

Изменение толщины переходного слоя, а следовательно, и емкости перехода при изменении смещения влияет на режим переменного тока транзистора, что важно для некоторых типов транзисторов и для некоторых областей рабочих частот. При изменении мгновенного напряжения на коллекторе транзистора с изменением напряжения усиливаемого сигнала меняется соответственно и мгновенная толщина барьера коллектора.

Поскольку часть коллекторного барьера находится с базовой стороны от границы база — коллектор, эффективная толщина базового слоя колеблется. Следовательно r_6 не является постоянным, как предполагалось до сих пор, а меняется в течение периода сигнала, и степень этого изменения зависит от толщины базового слоя, от того, на какое расстояние коллекторный барьер входит в него при нормальном коллекторном смещении, и от того, насколько велики колебания коллекторного напряжения. Далее, емкость коллектора C_k также меняется в течение периода сигнала. Очевидно, эти изменения надо принимать во внимание для полного анализа поведения транзистора. Влияние колебаний толщины коллекторного барьера может быть представлено включением дополнительных элементов в эквивалентную схему транзистора [Л. 2].

ЗАДАЧИ

1. Германиевый плоскостной диод (выпрямитель с *p-n* переходом) с площадью барьера 1 *мм*^{2} имеет сопротивления *p*- и *n*-областей соответственно 0,1 и 2 *ом · см* и время жизни неосновных носителей 100 и 200 *мксек* Покажите, что обратный ток утечки этого диода при комнатной температуре равен примерно 0,4 *мка.*

2. Докажите, что толщина барьера в диоде задачи 1, находящемся под обратным смещением 25 *в*, равна примерно $4 \cdot 10^{-4}$ *см*, а емкость — 25 *пф.*

3. Проверьте уравнение (19-11) для толщины барьера в диоде.

4. Покажите, что если *p-n* переход находится под прямым смещением, то сумма эффективностей эмиссии неосновных носителей в обоих направлениях равна единице.

5. Во многих случаях емкость барьера диода при постоянном токе имеет меньшее значение, чем дифференциальная емкость

dQ/dU_a. Докажите, что дифференциальная емкость симметричного p-n барьера равна

$$\frac{dQ}{dU_a} = \sqrt{\frac{EeN}{16\pi}} \times$$

$$\times \left(\frac{1}{\sqrt{U_0 - U_a}} + \frac{U_a}{2(U_0 - U_a)^{3/2}} \right).$$

ЛИТЕРАТУРА

1. Goucher F. S. et al., Theory and Experiment for a Germanium p-n Junction, Physical Review, 1951, vol. 81, p. 637.

2. Early J. M., Effect of Space-Charge Layer Widening in Junction Transistors, Proceedings of the Institute of Radio Engineers, 1952, vol. 40, p. 1401.

3. Миддлбрук Р. Д., Введение в теорию транзисторов. Атомиздат, 1960.

ГЛАВА ДВАДЦАТАЯ

ТЕОРИЯ КОНТАКТОВ МЕТАЛЛ—ПОЛУПРОВОДНИК

В предыдущей главе рассмотрены вопросы структуры и электрических свойств полупроводниковых p-n переходов и выведены уравнения зависимостей тока от напряжения, эффективности эмиссии неосновных носителей, толщины и емкости барьера. Проанализируем подобным же образом выпрямляющий контакт металла с полупроводником. Анализ будет таким же полным и будет основываться на установленных ранее положениях. В последующем следует иметь в виду, что если p-n переход имеет лишь один вид, то контакт металла с полупроводником может иметь две разновидности: металла с электронным полупроводником и металла с дырочным полупроводником. Анализ будет проводиться для контакта металла с электронным полупроводником, и лишь иногда параллельно будут даваться результаты для случая контакта с полупроводником p-типа.

20-1. РАСЧЕТ ЭЛЕКТРИЧЕСКОЙ ХАРАКТЕРИСТИКИ ИДЕАЛЬНОГО ВЫПРЯМИТЕЛЯ

Описание контакта металл — полупроводник начнем с верхней диаграммы на рис. 20-1, на которой изображены энергетические диаграммы для металла и типич-

ного электронного полупроводника. Предполагается, что оба вещества приготовлены в виде образцов с плоскими поверхностями. Представим себе, что плоская поверхность металла расположена близко к параллельной ей плоской поверхности полупроводника аналогично конденсатору с параллельными пластинами. Допустим, что первоначально оба образца были электрически нейтральными и пространство между ними было свободно от поля. Уровень Ферми в металле расположен у вер-

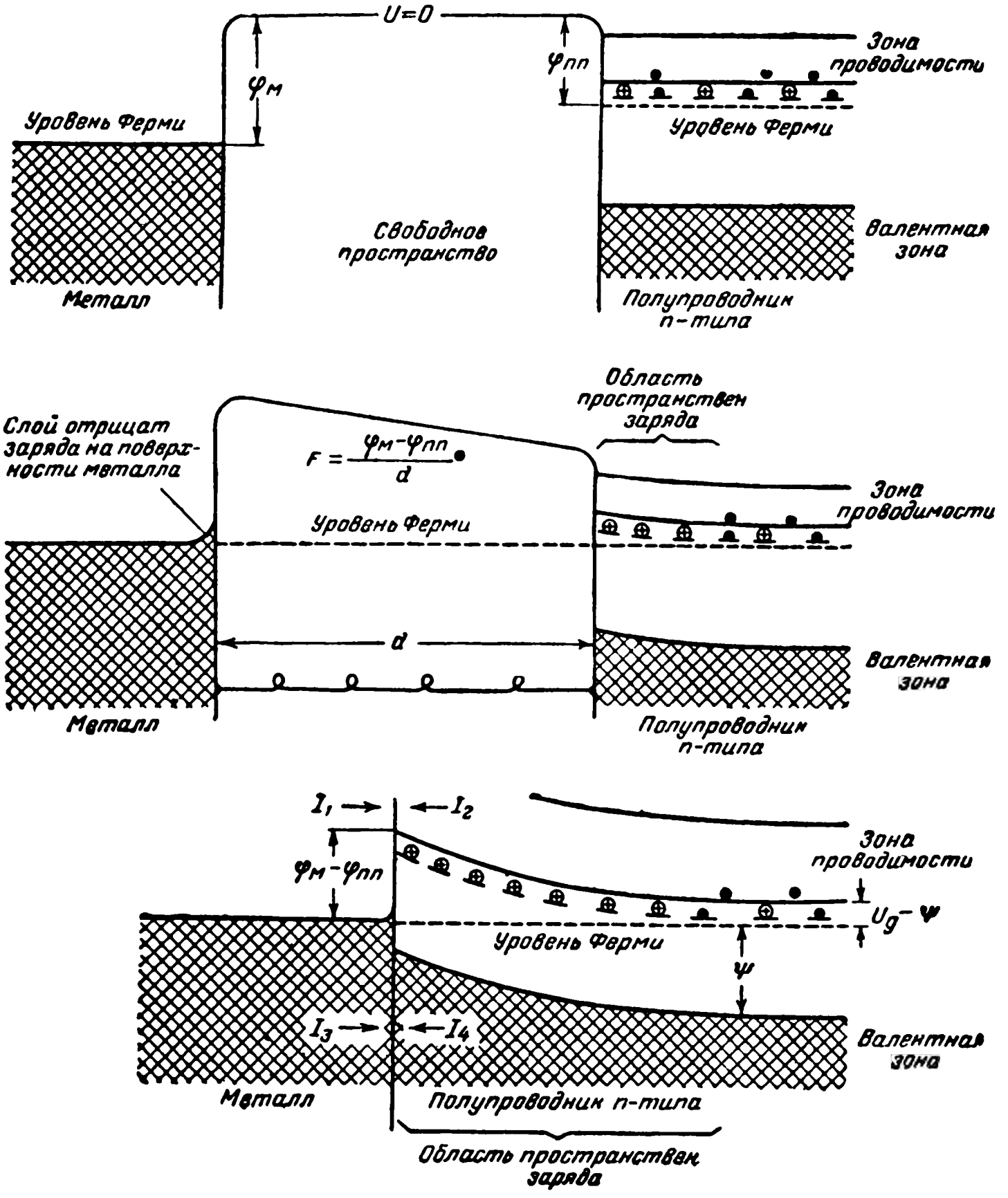


Рис. 20-1. Энергетические диаграммы для пар металл — полупроводник *n*-типа в трех случаях: образцы не контактируют (верхний рисунок), образцы соединены электрически (средний рисунок) и поверхности образцов соприкасающихся (нижних рисунок).

шины электронного распределения в частично заполненной валентной зоне. Верхняя граница этой зоны лежит выше уровня $U=0$, от которого ведется отсчет энергий (рис. 16-1 для энергетических зон типичных металлов — натрия и меди).

Работа выхода металла $\varphi_м$ есть разность (на графике по вертикали) потенциалов — соответствующего уровню Ферми внутри металла и потенциала пространства вне металла на большом расстоянии от его поверхности. Соответствующая величина для полупроводника обозначена $\varphi_{пп}$. На средней диаграмме на рис. 20-1 показано, что происходит, когда между металлом и полупроводником установлен электрический контакт посредством проволоки, соединяющей их поверхности. Вначале электроны перетекают из полупроводника в металл, пока последний не приобретет достаточный отрицательный заряд, а полупроводник — такой же положительный заряд, чтобы оба уровня Ферми сместились до совпадения при одной и той же энергии. При этом избыток отрицательного заряда в металле располагается в тончайшем слое на его поверхности, тогда как избыток положительного заряда в полупроводнике расположен в слое пространственного заряда довольно значительной толщины, состоящем из некомпенсированных ионизованных доноров, находящихся под поверхностью полупроводника. В пространстве между этими двумя поверхностями и возникает электрическое поле, причем такое, что интеграл его по расстоянию от одной до другой поверхности равен контактной разности потенциалов $\varphi_м - \varphi_{пп}$ между металлом и полупроводником.

Если эти две поверхности постепенно сближать, то электрическое поле между ними становится сильнее, емкость эквивалентного конденсатора увеличивается, величина заряда на поверхности металла и пространственный заряд внутри поверхностного слоя полупроводника возрастают, вследствие того что для поддержания постоянной разности потенциалов между пластинами конденсатора с увеличивающейся емкостью ему надо сообщать соответственно возрастающий заряд. Наконец, на нижней диаграмме на рис. 20-1 показан случай, когда поверхности приведены в соприкосновение без приложения внешнего напряжения. В этом случае поле, обусловленное контактной разностью потенциалов меж-

ду металлом и полупроводником, существует полностью внутри полупроводника направо от контакта. Внутри этой области пространственного заряда практически все доноры ионизованы, а бóльшая часть электронов проводимости в свободной зоне полупроводника вытолкнута из этой области действием электрического поля внутри нее. Чтобы пройти сквозь результирующий потенциальный барьер, электрон металла, либо полупроводника должен обладать энергией $e(\varphi_м - \varphi_{nn})$ сверх энергии Ферми.

Как в случае p-n перехода, электрическую характеристику выпрямителя с контактом металл — полупроводник можно вывести, разложив полный ток на четыре отдельные составляющие, и затем рассмотреть влияние приложенного напряжения на каждую составляющую в отдельности. К этому анализу мы сейчас и перейдем.

1. Составляющая I_1 создается потоком электронов из металла, обладающих достаточной энергией для преодоления барьера и движущихся в нужном направлении. Величина этого тока дается выражением, подобным уравнению Ричардсона для термоэлектронной эмиссии с поверхности металла:

$$I_1 = -QT^{3/2}e^{-e(\varphi_м - \varphi_{nn})/kT} \; a/см^2. \qquad (20\text{-}1)$$

Величина Q в этом выражении имеет меньшее числовое значение, чем величина тока насыщения по Ричардсону ($120\ a/см^2 \cdot град^{3/2}$) *, так как в твердом выпрямителе эмиссия происходит не в вакуум, а в близрасположенный полупроводник, откуда вследствие столкновений и отражений некоторые электроны возвращаются снова в металл.

2. Составляющая I_2 создается потоком электронов в зоне проводимоети полупроводника, движущихся к металлу и имеющих энергию, достаточную для преодоления барьера справа налево. При равновесии этот ток имеет ту же величину, что и I_1, но противоположен по направлению.

Отсюда

$$I_2 = QT^{3/2}e^{-e(\varphi_м - \varphi_{nn})/kT} \; a/см^2. \qquad (20\text{-}2)$$

3. Составляющая I_3 создается потоком положительных дырок из глубины заполненной электронами зоны металла в валентную зону полупроводника. Количественно она равна:

$$I_3 = \frac{eD_p}{\sqrt{D_p\tau_p}} 2UT^{3/2}\, e^{-e\Psi/kT} \; a/см^2, \qquad (20\text{-}3)$$

где Ψ имеет значение, показанное в нижней части рис. 20-1.

---

* Уравнение Ричардсона содержит множитель T^2, а не $T^{3/2}$ как в данном случае. Соответственно константа в уравнении Ричардсона равна $120\ a/см^2 \cdot град^2$ и отождествлять ее с Q нельзя. *Прим. ред.*

4. Составляющая I_4 создается потоком положительных дырок проходящих справа налево из валентной зоны полупроводника в металл. Если составляющие тока I_1, I_2 и I_3 можно рассматривать как термоэлектронные токи в том смысле, что они проходят над или под барьерами, ограничивающими ток, то составляющая I_4 возникает в результате непрерывной генерации новых положительных дырок на расстоянии диффузионной длины от барьера и благодаря непрерывной диффузии этих дырок в барьер и через него Величину I_4 можно получить подобно тому, как это сделано в гл. 19 для p-n перехода

$$I_4 = -\frac{eD_p}{\sqrt{D_p\tau_p}}\,2UT^{3/2}\,e^{-e\Psi/kT} \quad a/см^2. \tag{20-4}$$

При тепловом равновесии и отсутствии внешнего смещения на барьере составляющие I_1 и I_2 равны по величине и противоположны по знаку; подобным же образом I_3 и I_4 равны и противоположно направлены. Поэтому общий ток, проходящий через барьер, равен нулю.

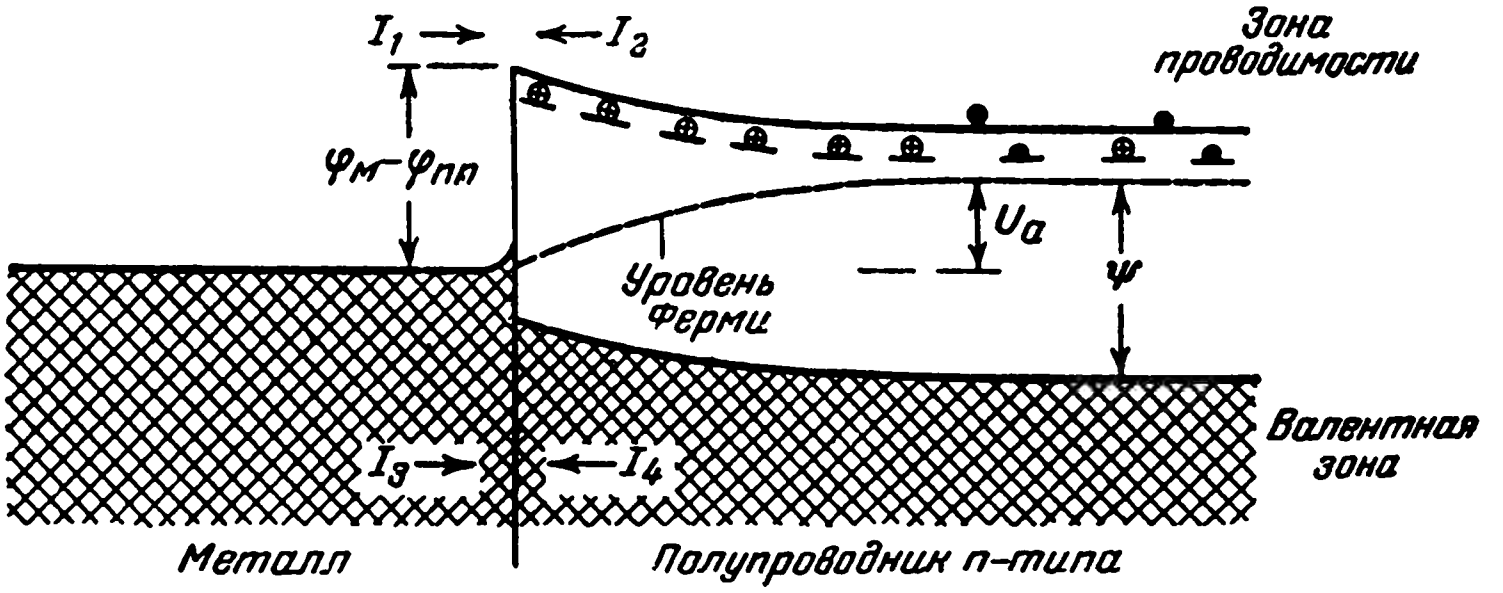


Рис 20-2. Составляющие тока I_2 и I_3 увеличиваются на множитель $e^{eU_a/kT}$, если к контакту, показанному на рис. 20-1, приложено прямое смещение.

Однако наложенное смещение нарушает это равновесие и вызывает увеличение составляющих I_2 и I_3. Нам далее и предстоит описать поведение получающегося в результате суммарного тока.

На рис. 20-2 показан этот выпрямляющий барьер при малом приложенном смещении U_a с полярностью: металл — плюс, полупроводник — минус. Это направление прямого смещения. Уровень Ферми уже не является уровнем постоянной во всей системе энергии, а испытывает смещение вверх на величину U_a в приконтактной области полупроводника. Область пространственного заряда барьера становится тоньше, а высота потенциаль-

ного барьера с точки зрения электрона, находящегося в полупроводнике и движущегося по направлению к металлу, уменьшается на величину U_a. Поскольку высота барьера для этих электронов сократилась, большее количество электронов может пересекать барьер справа налево и составляющая тока I_2 соответственно увеличивается в $e^{eU_a/kT}$ раз. Подобным же образом смещение перехода вызывает уменьшение высоты барьера,

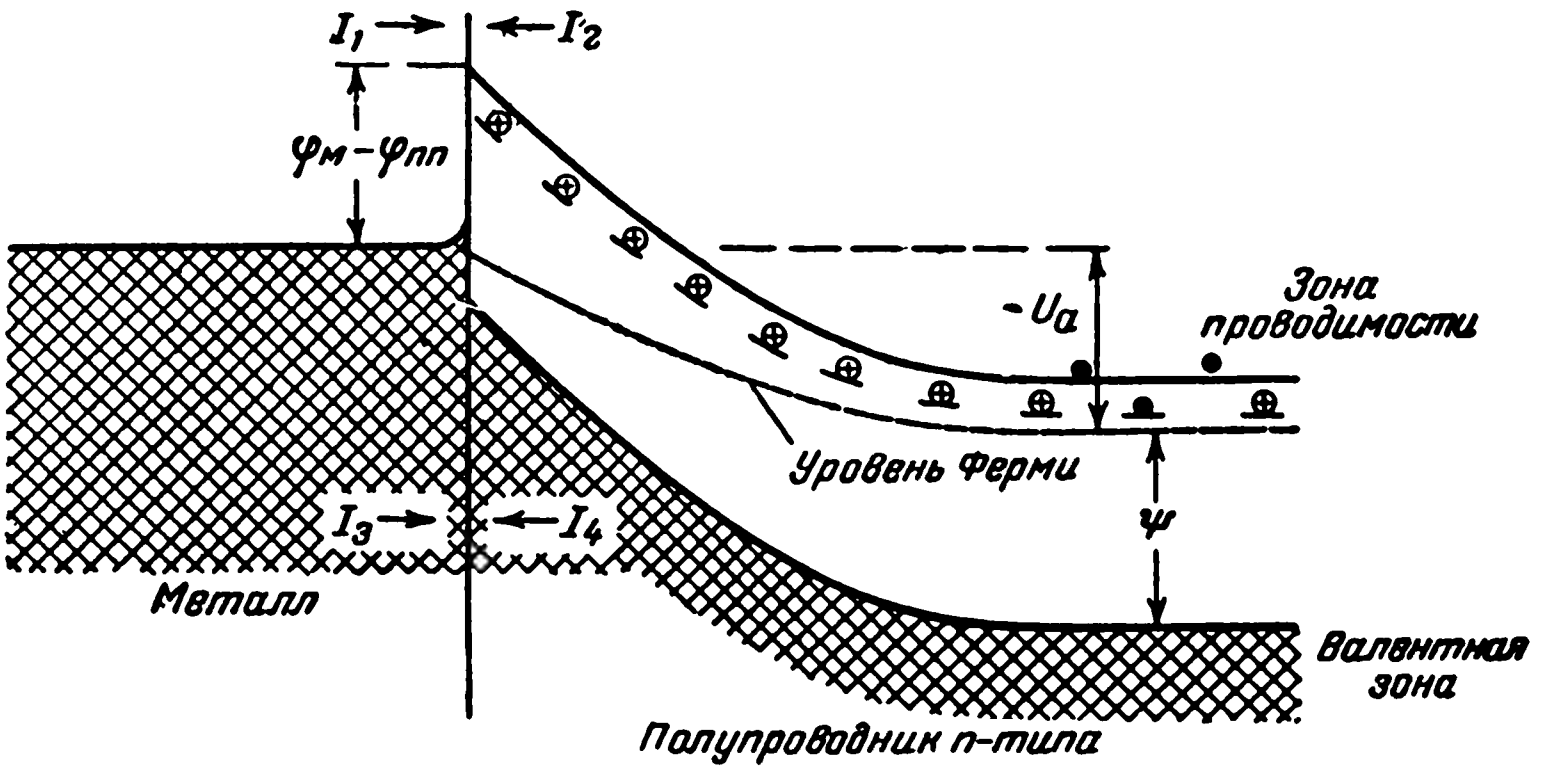


Рис. 20-3. Составляющие тока I_2 и I_3 уменьшаются в $e^{-eU_a/kT}$ раз по сравнению с их равновесным значением, если к контакту приложено обратное смещение.

который должны преодолеть положительные дырки, идущие из металла в полупроводник. Количество таких дырок увеличивается, причем тоже в $e^{eU_a/kT}$ раз, и составляющая I_3 увеличивается, следовательно, во столько же раз.

Однако такое прямое смещение барьера не меняет составляющих I_1 и I_4. Если величина прямого смещения больше 0,1 *в*, то суммарный ток в основном состоит из возросших составляющих I_2 и I_3, тогда как составляющие I_1 и I_4, хотя они и остаются постоянными, становятся незначительными сравнительно с I_2 и I_3.

На рис. 20-3 показан тот же барьер при малом обратном смещении с полярностью: металл — минус, полупроводник — плюс.

В этом случае равновесие, существовавшее при нулевом смещении, нарушается в другом направлении. Эффективная высота барьера, преодолеваемого электро-

нами при переходе из полупроводника в металл, увеличивается на величину приложенного смещения — U_a; следовательно, составляющая тока I_2 *уменьшается* в $e^{-eU_a/kT}$ раз по сравнению с ее равновесным значением и то же происходит с составляющей I_3. Обратное смещение не изменяет составляющих I_1 и I_4. При обратном смещении, превышающем 0,1 *в*, составляющие I_2 и I_3 становятся сравнительно незначительными и обратный ток делается равным току насыщения, как уже было показано в предыдущих главах.

Полный ток диода при любом смещении напряжения может быть представлен алгебраической суммой четырех составляющих, увеличенных или уменьшенных на соответствующие множители, зависящие от величины смещения.

Таким образом,

$$I_{\text{общ}} = I_1 + I_4 + (I_2 + I_3)\, e^{eU_a/kT} \quad a/см^2.$$

Так как

$$I_1 + I_4 = -(I_2 + I_3),$$

то

$$I_{\text{общ}} = (I_1 + I_4)(1 - e^{eU_a/kT}) = I_{\text{нас}}(1 - e^{eU_a/kT})\, a/см^2, \quad (20\text{-}5)$$

где U_a положительно, если смещение прямое, и отрицательно, если смещение обратное, и где

$$I_{\text{нас}} = I_1 + I_4 = -QT^{3/2} e^{(\varphi_м - \varphi_{nn})/kT} -$$

$$- \frac{eD_p}{\sqrt{D_p \tau_p}} \cdot 2UT^{3/2} e^{-e\Psi/kT} \quad a/см^2. \quad (20\text{-}6)$$

Таким образом, уравнение (20-5) приводит к той же электрической характеристике диода с контактом металл — полупроводник, что и зависимость, даваемая уравнением (19-9) для ***p-n*** **перехода.**

20-2. КОНТАКТ МЕТАЛЛА С ПОЛУПРОВОДНИКОМ *p*-ТИПА

Все рассмотренные в предыдущем параграфе положения в равной мере применимы и к контакту металла с полупроводником *p* - типа

В уравнениях необходимо лишь кое-где изменить индексы и поменять знаки.

На рис. 20-4 показаны энергетические диаграммы для контакта металла с полупроводником p-типа, для которых $\varphi_{пп} > \varphi_м$. В верхней части диаграммы показан барьер без приложенного смещения. В средней части

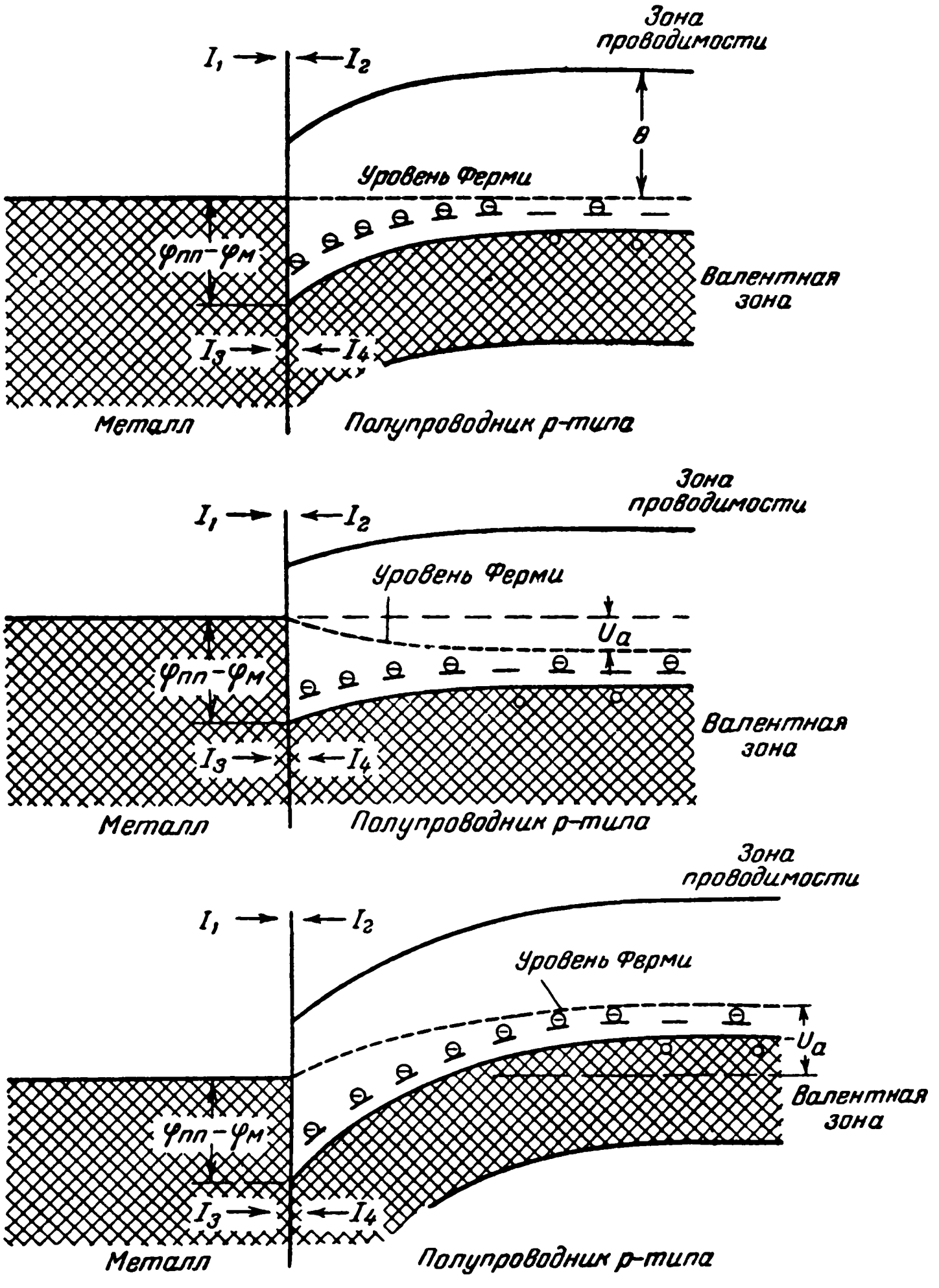


Рис 20-4 Энергетические диаграммы для контакта металла с полупроводником p-типа при нулевом смещении (верхний рисунок), прямом смещении (средний рисунок) и обратном смещении (нижний рисунок).

диаграммы показано влияние прямого смещения с полярностью: металл — минус, полупроводник — плюс. Прямое смещение способствует движению электронов из металла в зону проводимости полупроводника I_1, а дыркам — из валентной зоны полупроводника в металл I_4. В нижней части рис. 20-4 показано влияние обратного смещения, при котором число частиц в этих двух составляющих уменьшается. Если обратное смещение достаточно велико, то I_1 и I_4 становятся пренебрежимо малыми, а составляющие I_2 и I_3, не изменяющиеся под влиянием смещения, продолжают оставаться неизменными, составляя обратный ток насыщения I_2+I_3.

Следует отметить, что прямое направление смещения для диода p-типа противоположно прямому направлению для диода n-типа, о котором шла речь в предыдущем параграфе. Конечно, противоположны также и направления обратных смещений. Эта разница в направлениях выпрямления и является первым доказательством того, что существует два типа полупроводников. Исторически обозначения n и p были введены для того, чтобы различить полупроводники по этому их поведению, и эти обозначения первоначально относились к напряжениям смещения полупроводника для прямого направления тока через диод. Наблюдения этого различия побудили к теоретическим исследованиям, приведшим к выводу, что в полупроводниках p-типа ответственными за проводимость являются положительные дырки. Уравнения, определяющие четыре составляющие тока для диода с полупроводником p-типа, подобны уравнениям для диода с полупроводником n-типа и имеют вид:

$$I_1 = -\frac{eD_n}{\sqrt{D_n\tau_n}} \cdot 2UT^{3/2} e^{-e\theta/kT} \quad a/см^2; \tag{20-7}$$

$$I_2 = \frac{eD_n}{\sqrt{D_n\tau_n}} \cdot 2UT^{3/2} e^{-e\theta/kT} \quad a/см^2; \tag{20-8}$$

$$I_3 = QT^{3/2} e^{-e(\varphi_{nn}-\varphi_м)/kT} \quad a/см^2; \tag{20-9}$$

$$I_4 = -QT^{3/2} e^{-e(\varphi_{nn}-\varphi_м)/kT} \quad a/см^2; \tag{20-10}$$

$$I_{\text{нас}} = I_2 + I_3 =$$

$$= \frac{eD_n}{\sqrt{D_n \tau_n}} \cdot 2UT^{3/2} e^{-e\theta/kT} + QT^{3/2} e^{-e(\varphi_{nn} - \Phi_\varphi)/kT} \quad a/см^2; \tag{20-11}$$

$$I_{\text{общ}} = I_{\text{нас}} (1 - e^{eU_a/kT}) \; a/см^2, \tag{20-12}$$

где θ в уравнениях (20-7) и (20-8) имеет значение, показанное на рис. 20-4, и где, как и раньше, U_a принимается положительным для прямого смещения и отрицательным — для обратного. Следует отметить, что по уравнению (20-12) токи прямого направления имеют ог-

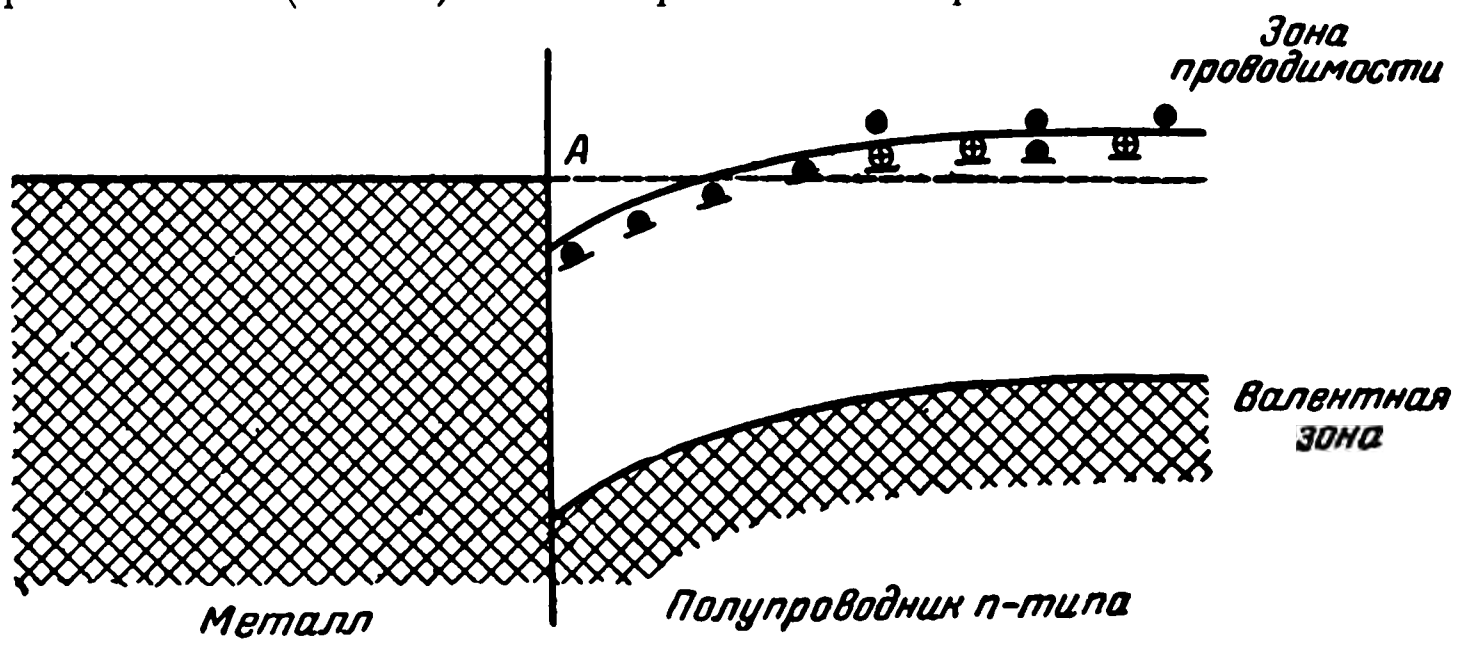


Рис 20-5 При контакте металла с полупроводником *n*-типа, у которого $\varphi_{nn} > \varphi_м$, непосредственно после установления механического контакта, но до достижения равновесия область *A* заполняется электронами и сопротивление контакта становится равным нулю.

рицательный знак. Этот результат математически объясняет тот факт, что прямой ток в выпрямителе *p*-типа идет в направлении, при котором положительный заряд перемещается справа налево, т. е. в отрицательном направлении, как это показано на диаграммах. В этом отношении диод *p*-типа отличается от диода *n*-типа, в котором прямые токи имеют положительный знак, что соответствует направлению условного положительного тока слева направо на этой диаграмме. Положения, рассмотренные в предыдущих двух параграфах, основаны на предположении, что $\varphi_м > \varphi_{nn}$ для контакта *n*-типа и что $\varphi_{nn} > \varphi_м$ для контакта *p*-типа. Если же для данного контакта не соблюдается соответствующее ему условие, барьера не будет; не будет и выпрямления. На рис. 20-5 показана энергетическая диаграмма для кон-

такта металл — полупроводник n-типа, для которого $\varphi_{nn} > \varphi_м$ и для которого поэтому нарушено первое из приведенных выше условий. На рисунке показано состояние, когда контакт уже осуществлен. Заметим, что вблизи металла нижняя граница зоны возбуждения полупроводника находится ниже уровня Ферми, что создает углубление A. Такое состояние динамически невозможно, и электроны немедленно потекут в это углубление, точно так же как вода станет заполнять углубление на поверхности жидкости. Это углубление в результате исчезает, и электроны получают возможность передвигаться беспрепятственно в любом направлении в зоне проводимости полупроводника. Те же доводы можно привести, чтобы показать, что в случае контакта металла с полупроводником p-типа при $\varphi_{nn} < \varphi_м$ барьера также не образуется.

Такой контакт идеален в смысле отсутствия сопротивления, и единственным сопротивлением в такой системе является сумма последовательных сопротивлений металла, тела полупроводника и остальных контактов присоединенной цепи. Невыпрямляющий контакт между металлом и полупроводником может быть получен, таким образом, если подобрать материалы так, что $\varphi_м < \varphi_{nn}$ для полупроводника n-типа или $\varphi_м > \varphi_{nn}$ для полупроводника p-типа. В практике создание хорошего невыпрямляющего контакта с малым сопротивлением определяется отчасти правильным подбором металла в соответствии со сказанным выше. Это сопротивление зависит также от физических условий, химической чистоты поверхностей и механической плотности контакта.

20-3. ВЫПРЯМИТЕЛЬНЫЕ РЯДЫ

Интересным подтверждением правильности теории выпрямления, изложенной в этой главе, является существование выпрямительных рядов, получаемых при измерении электрических характеристик контактов металлов с полупроводником, если выбран некоторый определенный полупроводник и ряд металлов с различными работами выхода. В ряде случаев было обнаружено, что вообще выпрямительные характеристики располагаются в том же порядке, что и работы выхода используемых металлов. Подобные ряды были экспериментально обнаружены в случаях контактов металлов

с окисью меди и селеном. Как окись меди, так и селен являются полупроводниками p-типа, и теоретическое обоснование этих закономерностей мы сделаем для контакта металла с полупроводником p-типа.

Важным параметром характеристики выпрямителя является сопротивление R_0 при нулевом смещении, которое на характеристике в логарифмических координатах представляется как сопротивление на участке кривой, идущей под углом 45° в области, где прямые и обратные ветви совпадают при очень малых напряжениях и токах. Это сопротивление нулевого смещения можно вычислить алгебраически с помощью дифференцирования уравнения (20-12), разрешая его относительно dU_a/dI и полагая $U_a \to 0$.

В результате получаем:

$$R_0 = -\frac{kT}{eI_{\text{нас}}} \; ом/см^2, \tag{20-13}$$

где на отрицательный знак можно не обращать внимания, так как он не имеет физического значения, являясь всего лишь результатом выбора нами положительного направления смещения. Для данного полупроводника это сопротивление нулевого смещения имеет различные значения при контакте с металлами, имеющими различные значения $\varphi_м$.

Эта зависимость становится очевидной, если объединить уравнения (20-13) и (20-11). Получим:

$$R_0 = \frac{kT}{e}\left\{\frac{1}{\dfrac{eD_n}{\sqrt{D_n\tau_n}} \cdot 2UT^{3/2}\, e^{-e\theta/kT} + QT^{3/2}\, e^{-e(\varphi_{nn}-\varphi_м)/kT}}\right\} ом/см^2. \tag{20-14}$$

Для значений $\varphi_м$, при которых $\varphi_{nn} - \varphi_м$ значительно меньше, чем θ*, первым членом знаменателя в уравнении (20-14) можно пренебречь. Тогда

$$R_0 \approx \frac{k}{120eT}\, e^{e(\varphi_{nn}-\varphi_м)/kT} \; ом/см^2. \tag{20-15}$$

---

* Если $\varphi_{nn} - \varphi_м$ больше, чем θ, то ток при нулевом смещении определяется главным образом составляющими I_1 и I_2, не зависящими от работы выхода металла.

Из этого уравнения видно, что для контактов между металлами и полупроводниками *p*-типа сопротивление нулевого смещения велико в случае металлов с малой работой выхода и уменьшается с переходом к металлам с большей работой выхода. Если работа выхода металла больше, чем работа выхода полупроводника, контакт перестает выпрямлять вообще и остается только сопротивление тела самого полупроводника.

Таково в действительности поведение семейства снятых экспериментально выпрямительных характеристик, представленных на рис. 20-5 для случая контакта различных металлов с селеном. Металлы напылялись на поверхность селена в вакууме; эта поверхность предварительно очищалась в том же вакууме путем нагрева до возгонки части селена. Здесь же приведены характеристики для контактов селена с такими металлами, как никель и золото, у которых, по-видимому, работа выхода больше, чем у селена, и которые соответственно не обладают выпрямляющими свойствами; их электрическая характеристика представляет собой прямую, наклоненную под углом 45°, являющуюся просто характеристикой сопротивления тела селена. При переходе к металлам со все более низким значением $\varphi_м$ сопротивление нулевого смещения соответствующих выпрямляющих пар увеличивается в соответствии с уравнением (20-15).

На рис. 20-7 показаны в логарифмическом масштабе полученные в этом опыте сопротивления нулевого смещения в зависимости от работы выхода контактирующего металла. Получается совершенно прямая линия, на которую не попадает только точка, соответствующая кадмию. Уравнение (20-15) предсказывает подобную линейную зависимость между $\lg R_0$ и $\varphi_м$, но дает величину наклона $-\frac{e}{2{,}303kT} = -16{,}8$, тогда как экспериментальная величина всего лишь —2,4!

Подобные же результаты по контактам металлов с селеном получены в [Л. 1]. Логарифмы «сопротивлений перехода» (не указано, каким способом проводится их измерение, но, по-видимому, при нулевом смещении), нанесенные на график в функции работы выхода металла, как на рис. 20-7, образуют прямую линию с наклоном — 2. Подобные же исследования с выпрямителями,

одним из элементов которых была окись меди, дали в общем те же результаты. Хотя эти результаты в ряде важных моментов и подтверждают теорию, они расходятся с нею в этом существенном вопросе, и до тех пор, пока мы не найдем объяснения такому несоответствию, мы

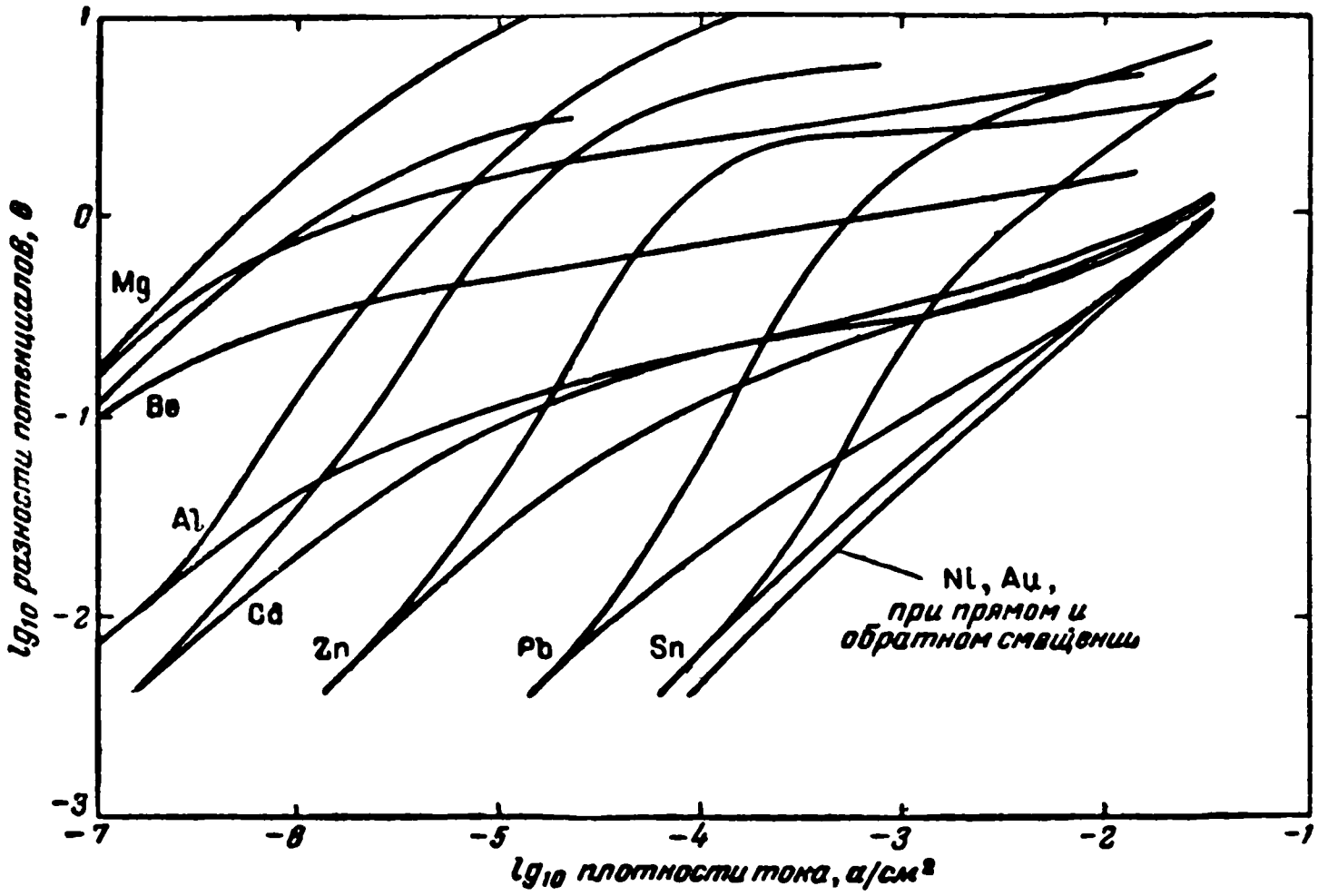


Рис 20-6 Смещение выпрямительных характеристик для контактов селена с различными металлами располагается в том же порядке, в котором расположены работы выхода металлов.

не можем утверждать, что знаем все о выпрямителях с контактом металла с полупроводником. Как показано в следующем параграфе, эффект притяжения электронов в полупроводнике их собственным электрическим отображением за поверхностью металла приводит к тому, что эффективная высота барьера становится меньше $\varphi_м - \varphi_{пп}$ для контактов n-типа и меньше $\varphi_{пп} - \varphi_м$ для контактов p-типа. К тому же, это уменьшение значительнее в случае барьеров большей высоты. Частично этим можно объяснить расхождение между теоретическими и экспериментальными значениями наклона прямой на рис. 20-7.

Пока нам недостаточно известно о влиянии таких не поддающихся учету факторов, как окисные пленки на контактирующих поверхностях, химические слои, ко-

торые образуются у контактов, и поверхностные энергетические уровни [Л. 2—4], мы можем отнести все оставшиеся расхождения за счет этих влияний.

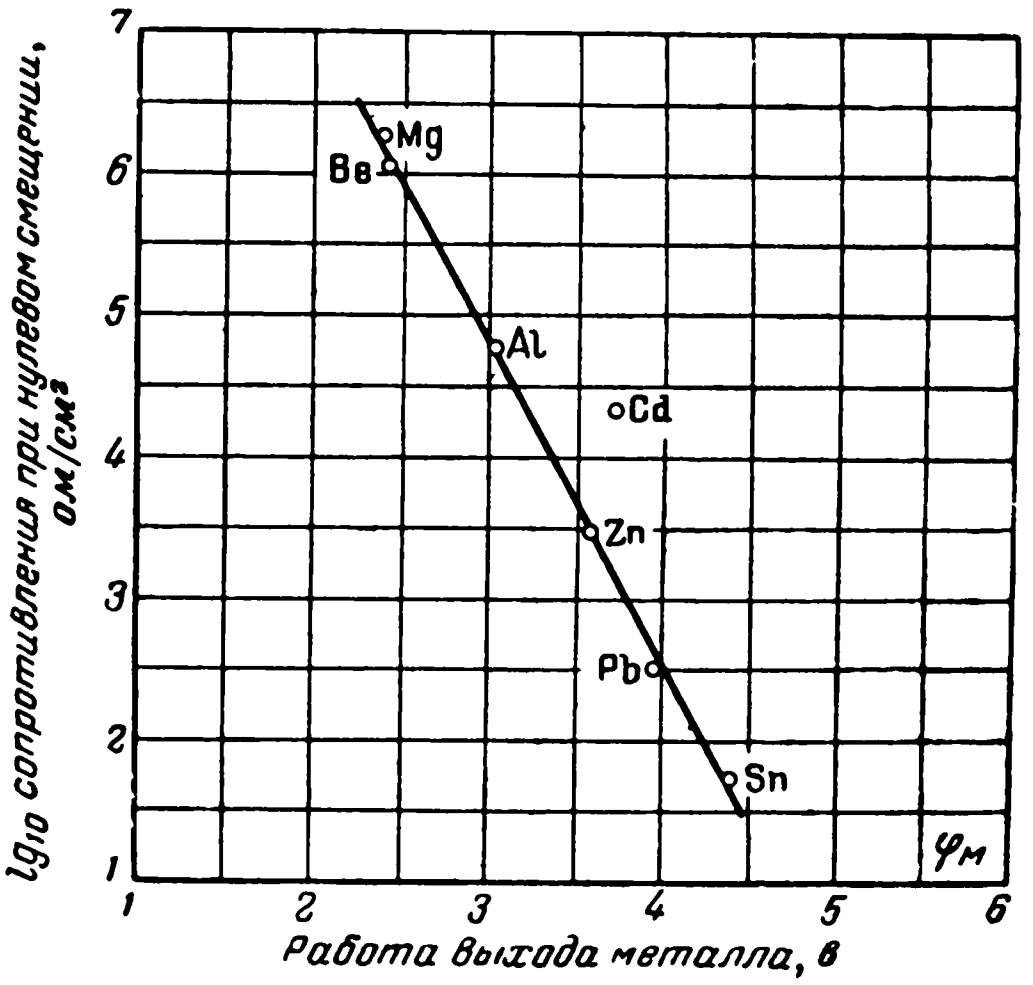


Рис 20-7 Сопротивление контакта металла с полупроводником при нулевом смещении экспоненциально зависит от работы выхода металла Представлены результаты для селена и различных металлов по рис 20-6.
Ni и Au не выпрямляют, и величина сопротивления их барьеров при нулевом смещении ничтожна сравнительно с величиной сопротивления полупроводника

20-4. ВЛИЯНИЕ ПОЛЯ ЭЛЕКТРИЧЕСКОГО ИЗОБРАЖЕНИЯ НА ХАРАКТЕРИСТИКУ ПЕРЕХОДА

Уравнения (20-5) и (20-12) дают для выпрямляющего металлополупроводникового контакта такую же электрическую характеристику, какая выведена нами в предыдущей главе для *p-n* перехода. Справедливость этих характеристик ограничена, конечно, необходимостью учитывать влияние сопротивления тела самого полупроводника при больших прямых токах и пробой барьера при высоких обратных напряжениях.

При рассмотрении экспериментальных электрических характеристик контактных выпрямителей (см., например, рис. 5-3, 5-5 и 5-12) всегда можно наблюдать откло-

нения от насыщения, предсказываемого теоретически для идеальной характеристики обратного тока и имеющего место в действительности в выпрямителях с p-n переходом (рис. 5-15). Частью это расхождение можно приписать омической (или близкой к омической) утечке по краям выпрямляющего контакта. Однако бóльшая часть ее, по-видимому, обусловлена изменением высоты барьера под действием приложенного смещения [Л. 5 и 6].

Чтобы понять, как осуществляется эта зависимость, например, в случае контакта металла с полупроводником n-типа, рассмотрим электрон вне металла, но на расстоянии в несколько атомных диаметров от поверхности. Присутствие электрона индуцирует положительный заряд на поверхности металла, и электрон притягивается этим положительным зарядом.

Потенциал вблизи поверхности уменьшается действием этого индуцированного «изображения заряда», и для того чтобы удалить электрон, необходимо совершить работу. В самом деле, бóльшая часть термоэлектронной работы выхода металлической поверхности, по-видимому, состоит просто из работы, необходимой для того, чтобы удалить электрон на достаточно большое расстояние против силы притяжения его собственного индуцированного изображения.

Уменьшение потенциальной энергии электрона вблизи поверхности металла приводит к искривлению энергетических зон в полупроводнике около границы с металлом. Это искривление энергетических зон показано на рис. 20-8 для нижнего края зоны проводимости полупроводника. На этом рисунке показано, что эффективная высота барьера для составляющих тока I_1 и I_2 равна не $\varphi_м - \varphi_{пп}$ как предполагалось ранее, а $\varphi_м - \varphi_{пп} - \delta$, где δ — малая величина. По мере того как электрическое поле у поверхности металла возрасгает под действием приложенного обратного смещения, величина δ продолжает увеличиваться, уменьшая *эффективную* высоту барьера и увеличивая составляющую I_1 обратного тока. Эта картина подобна по своему физическому происхождению явлению при термоэлектронной эмиссии, когда эффективная работа выхода эмитирующего катода уменьшается при возникновении электрического поля у поверхности катода, что приводит к усилению эмиссии. Это поле можно настолько увеличить, что электроны

эмитируют уже без нагревания катода и имеет место ток электростатической электронной эмиссии. Поля такой величины, которая необходима для существенного понижения высоты барьера, действуют при обратном смещении в барьерах контактных выпрямителей, и соответствующие обратные токи уклоняются от идеального

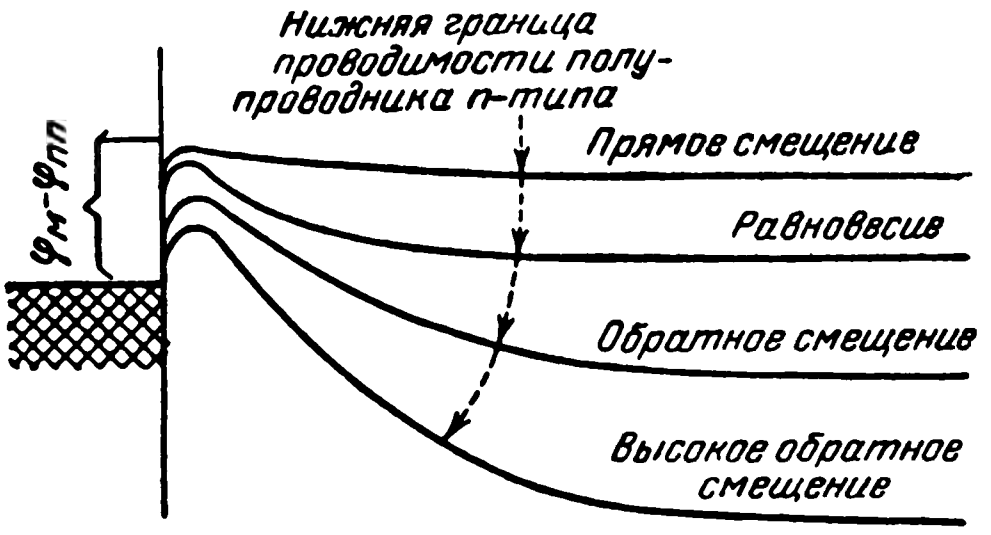


Рис 20-8 В контакте металла с полупроводником эффективная высота барьера меньше, чем $\varphi_м - \varphi_{пп}$, на величину зависящую от смещения.

насыщения с увеличением обратного смещения сначала медленно, а затем все более быстро, по мере того как постепенно понижается высота барьера. Даже и при прямом смещении эффективная высота барьера оказывается несколько пониженной. С увеличением прямого смещения δ становится меньше и высота барьера приближается к идеальной величине $\varphi_м - \varphi_{пп}$.

Величину δ можно без труда вычислить с помощью элементарной электростатики.

В результате получим:

$$\delta = \left(\frac{2e}{\bullet}\right)^{3/4} \{\pi N_d (\varphi_м - \varphi_{пп} - V_g + \Psi - U_a)\}^{1/4}. \quad (20\text{-}16)$$

где все величины выражены в электростатических единицах, или

$$\delta = 1{,}66 \cdot 10^{-5} \left(\frac{1}{\bullet}\right)^{3/4} \{N_d (\varphi_м - \varphi_{пп} - V_g + \Psi - U_a)\}^{1/4} \text{ см}, \quad (20\text{-}17)$$

где все величины даны в практических единицах и где, как и прежде, U_a положительно для прямого смещения и отрицательно для обратного смещения.

Влияние этой силы электрического изображения на все характеристики выпрямителей можно определить, заменив $\varphi_м - \varphi_{пп}$ в уравнениях для выпрямителя n-типа выражением $\varphi_м - \varphi_{пп} - \delta$, взяв δ из уравнения (20-17). На рис. 20-9 показана видоизмененная характеристика, рассчитанная для гипотетического выпрямителя n-типа с параметрами, данными под рисунком. Для сравнения показана также характеристика (пунктирной линией) того же выпрямителя, рассчитанная для идеального случая без учета поправки на влияние электрического изображения. Даже без приложенного внешнего смещения всегда присутствует собственное поле барьера и его влияние на понижение высоты барьера для составляющих I_1 и I_2 видно по сдвигу области нулевого смещения характеристики в сторону бóльших токов. В этом частном случае сдвиг равен примерно одному порядку, что соответствует уменьшению высоты барьера при нулевом смещении примерно на $^1/_{15}$ *в*.

Когда действует прямое смещение, δ уменьшается, высота барьера приближается к идеальной величине $\varphi_м - \varphi_{пп}$ и при большом прямом смещении сплошная кривая приближается к пунктирной. При обратном направлении смещения δ постепенно увеличивается с ростом обратного смещения и характеристика уклоняется ко все бóльшим обратным токам, все более отклоняясь от идеального тока насыщения.

То обстоятельство, что сплошная кривая на рис. 20-9 подобна по форме наблюдаемым характеристикам в большинстве выпрямителей с контактом металла с полупроводником, подтверждает теоретические выводы и объясняет многие характерные свойства выпрямителей, с которыми приходится сталкиваться на практике.

Подобный анализ, выполненный для случая контакта металла с полупроводником p-типа, приводит к выражению, сходному с уравнением (20-17):

$$\delta_p = 1{,}66 \cdot 10^{-5} \left(\frac{1}{\varepsilon}\right)^{3/4} \{N_a (\varphi_{пп} - \varphi_м - V_g + \theta - U_a)\}^{1/4} \text{ в}, \tag{20-18}$$

где U_a надо считать положительным для прямого смещения.

Эффективная высота барьера в этом случае $\varphi_{пп} - \varphi_м - \delta$, влияние величины δ на электрические характеристики прямого и обратного токов то же, что и только что описанное для контакта металла с полупроводником n-типа.

Все еще открытым остается вопрос относительно возможности возникновения лавинного пробоя в контактах

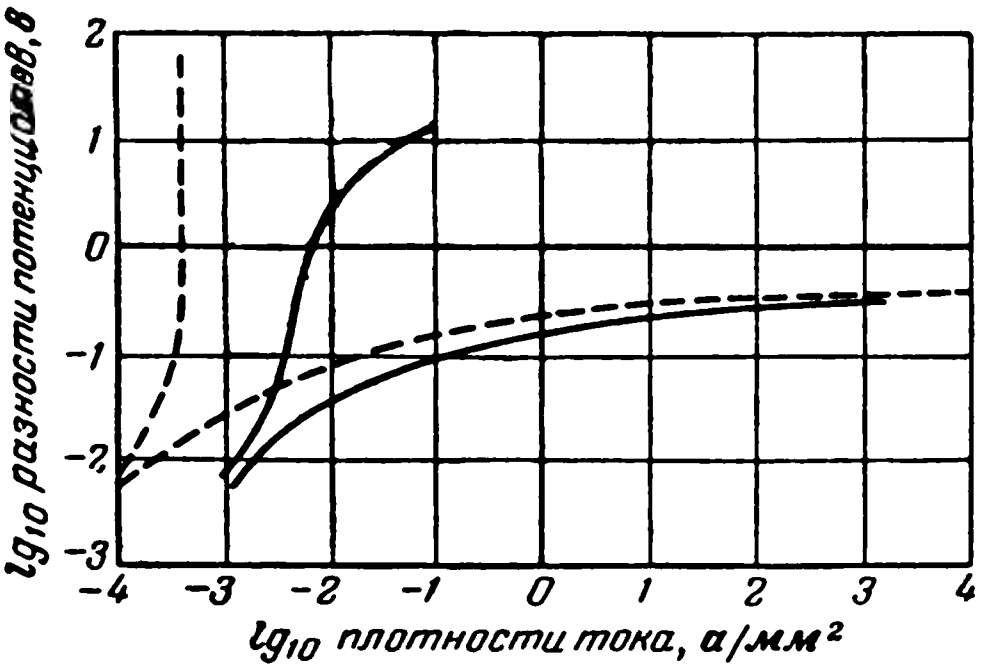


Рис. 20-9 Характеристика выпрямителя (сплошная линия), учитывающая уменьшение эффективной высоты барьера под действием сил электрического изображения, отличается от простейшей идеальной характеристики (пунктир).
Для расчета приняты следующие величины $\varphi_м - \varphi_{пп} = 0{,}5$ *в*; $\tau_p = 100$ *мксек*, $\Psi = 0{,}6$ *в*, $N_d = 10^{18}$ $см^{-3}$; $V_g = 0{,}72$ *в*; $T = 300°$ K; $D_p = 44$ $см^2/сек$.

металла с полупроводником. У полупроводников, имеющих высокую концентрацию доноров или акцепторов, толщина барьера сравнительно мала, и при больших обратных напряжениях безусловно создаются очень высокие поля, что может привести к пробою. Однако для таких барьеров δ тоже сравнительно велико и его дальнейшее увеличение при обратном смещении, по-видимому, достаточно хорошо объясняет форму обратной характеристики и маскирует могущие иметь место лавинные эффекты. Для большинства контактов металла с полупроводником точные количественные расчеты теоретической формы характеристики практически невозможны вследствие отсутствия в настоящее время точных данных о значениях таких величин, как $\varphi_{пп} - \varphi_м$ Ψ или δ

и N_d или N_a. Поэтому нельзя определенно сказать, что влияние сил электрического изображения, понижающих барьер, всегда предшествует истинному лавинному пробою. До настоящего времени не опубликовано материалов по наблюдениям лавинного пробоя барьеров в контактах металла с полупроводником.

20-5. ТОЧЕЧНЫЕ КОНТАКТЫ КАК ЭМИТТЕРЫ НЕОСНОВНЫХ НОСИТЕЛЕЙ

Остановимся теперь на использовании контактов металла с полупроводником в качестве эмиттеров неосновных носителей в полупроводнике. В случае контакта металла с полупроводником n-типа (рис. 20-2) составляющая прямого тока I_2 представляет собой ток основных носителей из полупроводника, а I_3 — эмиссионный ток неосновных носителей в полупроводнике. Отношение I_3 к I_2+I_3, таким образом, дает эффективность эмиссии неосновных носителей γ. Из рисунка можно видеть, что I_3 тем больше по сравнению с I_2+I_3, чем больше $\varphi_м-\varphi_{nn}$ или меньше ψ. Первое условие достигается выбором металла с высокой работой выхода, второе — созданием незначительной концентрации примесей в полупроводнике, чтобы уровень Ферми лежал значительно ниже нижней границы зоны проводимости.

В случае контакта металла с проводником p-типа для получения высокой эффективности эмиссии неосновных носителей следует подбирать металл с низкой работой выхода и обеспечивать малую концентрацию примесей в полупроводнике.

Как уже было отмечено, теоретические соображения, приведенные в этой главе относительно выпрямителей с контактом металла с полупроводником, объясняют их поведение лишь частично. Выпрямители не всегда и даже редко ведут себя в деталях в полном согласии с предсказаниями простой теории. Выпрямление в значительной степени зависит от технологии подготовки поверхностей до получения контакта с металлом, а теория, приведенная выше, вовсе не учитывает того, что на этих поверхностях могут образоваться химические слои. адсорбированные пленки, различные загрязнения и т. п. Эффективность эмиссии неосновных носителей зависит также от поверхностных условий, и получение хороших

контактных эмиттеров является в настоящее время примером в такой же мере результатов технологических проб и опыта, как и теоретического предвидения.

ЗАДАЧИ

1. Вычертите энергетическую диаграмму для контакта металла с полупроводником p-типа в случае $\varphi_{nn} < \varphi_м$ и покажите, что такой контакт не имеет сопротивления.

2. Влияет ли (и если да, то как) понижение барьера полем электрического изображения на емкость перехода металл—полупроводник?

3. Опишите, какие можно ожидать выпрямительные ряды для последовательностей переходов между двумя областями одного и того же полупроводника, имеющего следующие комбинации активирования: p^{++}-p^{+}, p^{++}-p, p^{++}-i, p^{++}-n, p^{++}-n^{+}, p^{++}-n^{++}. Вычертите качественно график, подобный графику на рис. 20-7, показывающий зависимость сопротивления нулевого смещения от контактной разности потенциалов перехода. Какая особенность наблюдается в случае p^{++}-i?

ЛИТЕРАТУРА

1. Schottky W., Abweichungen vom Ohmschen Gesetz in Halbleitern, Physikalische Zeitschrift, 1940, vol. 41, p. 570.

2. Tamm Ig., Über eine mögliche Art der Elektronenbindung an Kristalloberflächen, Physikalische Zeitschrift der Sowjetunion, 1932, vol. 1, p. 733.

3. Bardeen J., Surface States and Rectification at a Metal-Semiconductor Contact, Physical Review, 1947, vol. 71, p. 717.

4. Brattain W. H., Shockley W., Density of Surface States on Silicon Deduced from Contact Potential Measurements, Physical Review, 1947, vol. 72, p. 345.

5. Bethe H. A., NDRC Div 14 Report, MIT Radiation Laboratory No. 43—12, Nov. 23, 1942.

6. Meyerhof W. E., Contact Potential Difference in Silicon Crystal Rectifiers, Physical Review, 1947, vol. 71, p. 727.

7. Бонч-Бруевич В. Л., Теория полупроводников на VIII Всесоюзной конференции по полупроводникам, УФН, 1956, 60, 2, 213—224.

ГЛАВА ДВАДЦАТЬ ПЕРВАЯ

ТЕОРИЯ ПЛОСКОСТНОГО ТРИОДА

Рассмотрев в гл. 19 теорию p-n перехода, мы можем перейти теперь к описанию плоскостного триода. Первыми транзисторами были транзисторы с точечным контактом, но для разработки количественной теории этих приборов более подходят плоскостные n-p-n или p-n-p структуры с одномерным полем токов между параллель-

ными переходами; в дальнейшем мы будем всегда рассматривать только такую геометрию. В действительности нового сказать придется немного, так как плоскостной триод — это просто комбинация из двух параллельных диодных переходов, расположенных в пределах диффузионной длины неосновных носителей, и его работа определяется соответствующими смещениями, входным сигналом и выходной нагрузкой. Зная, как рабо-

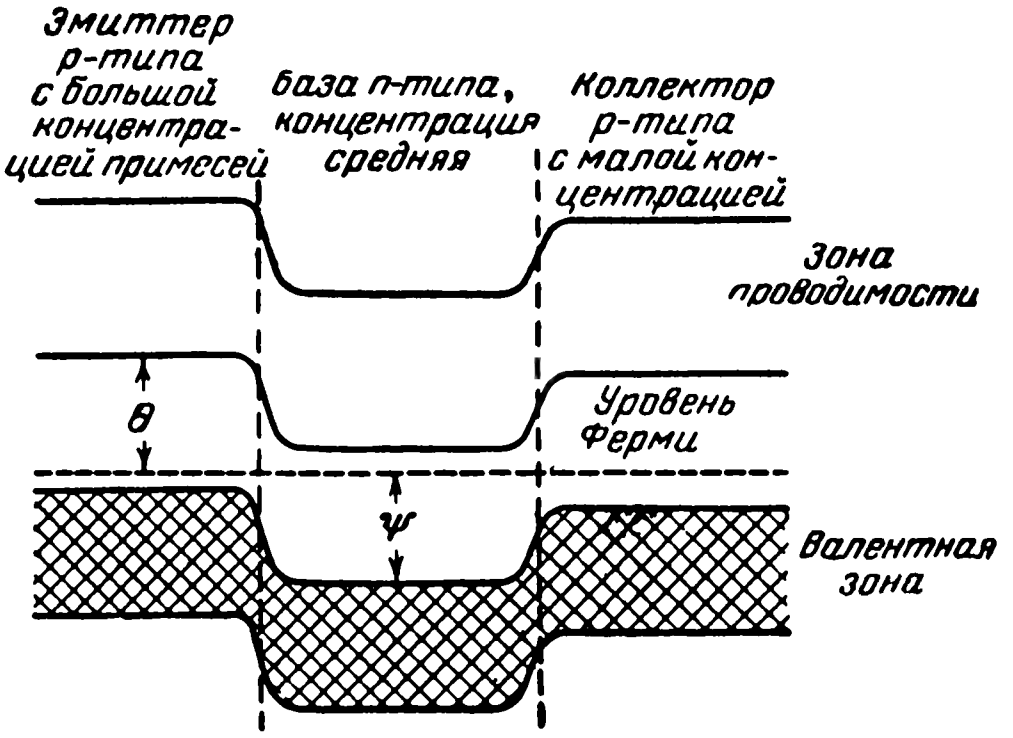


Рис. 21-1. Энергетическая диаграмма транзистора с *p-n-p* структурой в равновесии при нулевом смещении.

тает диод, нетрудно понять принцип действия триода. Однако в уравнения, выведенные в гл. 19 для диода, придется ввести некоторые количественные изменения. Эти изменения должны учитывать нарушения в поведении уединенного перехода, происходящие в случае помещения вблизи него другого перехода при обратном или прямом смещении.

В предыдущих главах работа триода описывалась как инъекция в базу (область дрейфа) неосновных носителей, концентрация которых колеблется в соответствии с сигналом, диффузия их через базовый слой и собирание их коллектором с другой стороны от базы в условиях, позволяющих получить существенный выигрыш в мощности. Как все это осуществляется в случае *n-p-n* триода, показано на энергетических диаграммах на рис. 21-1 и 21-2.

На рис. 21-1 переходы показаны в состоянии равновесия при нулевом смещении. База и коллектор этого

прибора имеют сравнительно небольшую концентрацию примесей, а эмиттер — большую концентрацию, с тем чтобы эффективность эмиссии положительных дырок в базу была достаточно высокой (рис. 21-1).

Малая концентрация примесей в коллекторе нужна, кроме того, чтобы область пространственного заряда коллекторного барьера оставалась относительно широ-

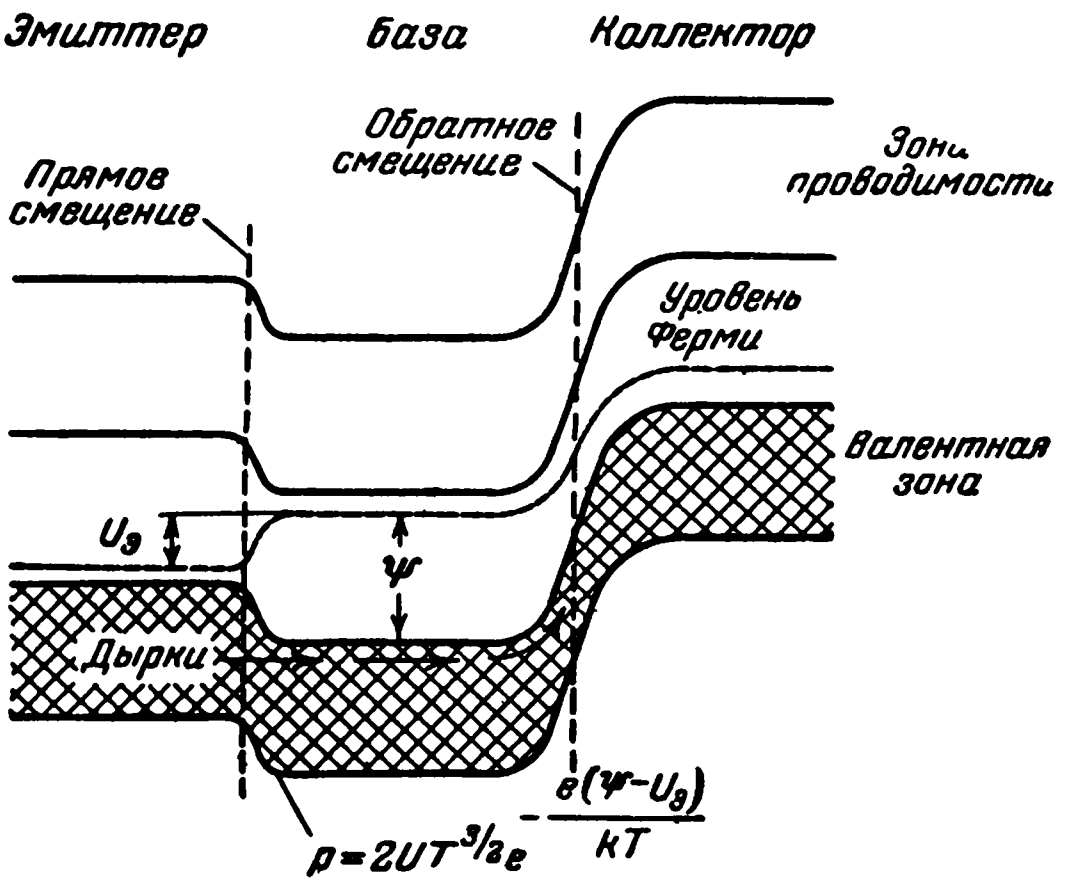


Рис. 21-2. Транзистор с *p-n-p* структурой (тот же, что на рис. 21-1) при прямом смещении на эмиттере и обратном смещении на коллекторе.

кой. В этом случае на коллекторный барьер можно накладывать высокие обратные смещения, не превышая поля пробоя внутри него, и, следовательно, возможны большие изменения коллекторного напряжения при работе. Малая концентрация примесей в базовом слое подобным же образом приводит к достаточно большой толщине коллекторного барьера, но, с другой стороны, она вызывает и нежелательный эффект: $r_б$ становится большим, и это приводит к понижению граничной частоты (§ 12-2). Таким образом, необходим какой-то компромисс в вопросе о концентрации примесей в базовом слое. При изготовлении выращенных германиевых транзисторов концентрация примесей в коллекторе может соответствовать удельному сопротивлению коллектора 10—20 *ом·см* при удельном сопротивлении базового

слоя около 1 *ом · см* и эмитерного слоя порядка 0,001—0,01 *ом · см.*

На рис. 21-2 показана комбинация переходов со смещениями, соответствующими нормальной работе транзистора как усилителя с общей базой. К эмитерному переходу приложено прямое смещение 0,2—0,3 *в*, а к коллекторному переходу — обратное смещение в несколько единиц или десятков вольт. При таких смещениях под эмитерным барьером в валентной зоне течет постоянный ток положительных дырок, диффундирующих через базовый слой с небольшой потерей в числе за счет рекомбинации; далее дырки идут на диаграмме вверх под коллекторным барьером в коллекторную область. Колебания эмитерного тока, являющиеся результатом наложения переменного сигнала на эмитерное постоянное смещение, вызывают соответствующие колебания тока неосновных носителей, поступающих в коллектор.

Физические параметры, важные для этого случая, — это ток эмиттера при прямом смещении, эффективность эмиссии неосновных носителей, входное дифференциальное сопротивление эмиттера, эффективность переноса частиц от эмиттера к коллектору (коэффициент передачи тока), коллекторное усиление и обратный ток насыщения коллектора. При работе на высоких частотах приходится учитывать также зависимость этих параметров от частоты. Ниже мы рассмотрим отдельно каждый из этих параметров, покажем, как они вытекают из физических концепций, разобранных в предыдущих главах, и как эти параметры должны быть взаимосвязаны, чтобы дать полную наблюдаемую картину работы транзистора.

21-1. ПРЯМОЙ ТОК ЭМИТТЕРА И ЭФФЕКТИВНОСТЬ ЭМИССИИ НЕОСНОВНЫХ НОСИТЕЛЕЙ

В гл. 19 рассмотрена эффективность эмиссии неосновных носителей *p-n* переходов. Указывалось на то, что очень важно иметь эффективный эмиттер, приводились соображения относительно концентрации примесей, от которых зависит эффективность эмитерных переходов. Было дано выражение для эффективности *p-n* перехода.

Положения, развитые в этой главе о прямом токе и эффективности эмиссии для уединенного *p-n* перехо-

да, надо для приложения к случаю эмиттера плоскостного триода несколько видоизменить. Необходимость эта вызвана тем, что на эти параметры оказывает влияние присутствие поблизости коллекторного перехода с обратным смещением. Это влияние на прямой ток можно показать следующим образом. Коллекторный переход при обратном смещении действует как неограниченный сток для неосновных частиц, диффундирующих к нему

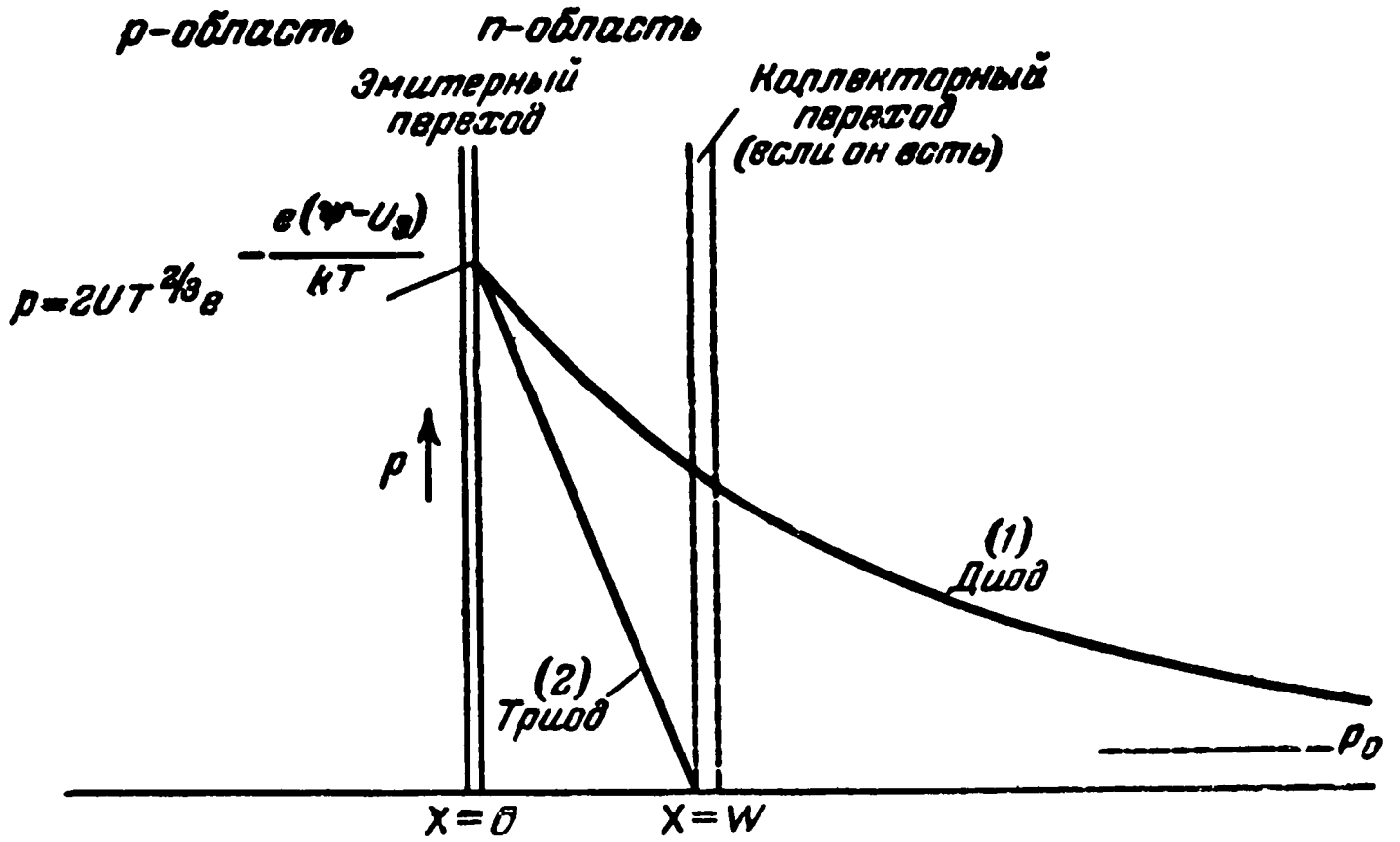


Рис. 21-3. Концентрация неосновных носителей (дырок) в *n*-области *p-n* перехода при прямом смещении зависит от наличия или отсутствия поблизости коллекторного перехода с обратным смещением.

через базовый слой. Такой коллектор, расположенный близко к эмиттеру, находящемуся при прямом смещении, ускоряет, таким образом, удаление неосновных носителей с базовой стороны эмитерного перехода сравнительно с диффузионным удалением их при отсутствии коллекторного перехода. Это приводит к увеличению тока неосновных носителей через базовый слой и, следовательно, к увеличению потока неосновных носителей через эмитерный переход по сравнению с величиной соответствующего тока по уравнению для диода при прямом смещении.

Нетрудно аналитически описать результат с помощью следующих рассуждений, основанных на рассмотрении графика градиента концентрации носителей, показанного на рис. 21-3. Слева на этом графике расположен *p-n* переход с большой концентрацией примесей в *p*-области

и малой концентрацией примесей в n-области при прямом смещении напряжения $U_э$. Этот переход может быть либо уединенным диодным переходом, либо эмитерным переходом p-n-p триода с коллектором при обратном смещении, показанным на рисунке пунктиром. На вертикальной оси отложена концентрация эмитируемых положительных дырок справа от первого p-n перехода, а две сплошные линии на графике дают концентрацию положительных дырок в зависимости от расстояния: 1) для плоскостного диода, описанного в гл. 19, и 2) для p-n-p триода. Следует отметить, что в случае диода концентрация дырок уменьшается экспоненциально до p_0 с постоянной спада L_p, как показано в решении примера 3 в гл. 18.

С другой стороны, для p-n-p триода, поскольку коллекторный переход при обратном смещении действует как неограниченный сток для положительных дырок, эмитируемых в область базы, концентрация дырок и градиент концентрации должны приводить, как показано на рисунке, к $p=0$ при $x=l$. Если $l \ll L_p$, этот градиент будет почти постоянным по всей толщине базового слоя. Ясно, что для данного значения p_1 концентрации положительных дырок на правой границе эмитерного перехода ($x=0$) ток сквозь переход в случае триода больше, чем в случае диода, так как градиент dp/dx при $x=0$ больше для триода из-за присутствия коллектора. Насколько этот ток больше, нетрудно показать вычислением токов для диода и триода.

Из примера 3 гл. 18

$$p - p_0 = (p_1 - p_0)\, e^{-x/L_p}.$$

Поэтому градиент p равен:

$$\frac{dp}{dx} = -\left(\frac{p - p_0}{L_p}\right) e^{-x/L_p}.$$

Определяя этот градиент при $x=0$, находим:

$$\left.\frac{dp}{dx}\right|_{x=0} (\text{диода}) = \frac{p_1 - p_0}{L_p} \approx -\frac{p_1}{L_p}; \quad \left.\frac{dp}{dx}\right|_{x=0} (\text{триода}) \approx -\frac{p_1}{l}.$$

Сравнение этих двух градиентов, которые в обоих случаях прямо пропорциональны прямым токам перехода, показывает, что при данном значении p_1 составляю-

щая I_3 прямого тока эмитерного перехода в L_p/l раз больше для триода, чем для диода с уединенным переходом. В этих двух случаях p_1 имеет одну и ту же величину для одинаковых эмитерных смещений, так как концентрация неосновных носителей с правой стороны эмитерного перехода в обоих случаях равна

$$2UT^{3/2}e^{-e(\Psi-U_э)/kT}$$

и пополняется из огромного количества дырок в p-области перехода, по мере того как дырки удаляются вправо под влиянием градиента концентрации в n-области.

Так как величина составляющей неосновных носителей прямого эмитерного тока возрастает из-за наличия коллекторного перехода с обратным смещением, очевидно, что эффективность его эмиссии ближе к единице, чем в случае диода в соответствии с этим же множителем.

Уравнение эффективности эмиттера для триода принимает в результате вид:

$$\gamma = \frac{1}{1+\frac{\sigma_n l}{\sigma_p L_n}}. \qquad (21\text{-}1)$$

Поэтому для получения высокой эффективности эмиттера желательно иметь большое различие между проводимостью эмиттера и базы, тонкий базовый слой и большое время жизни электронов в эмиттере. К сожалению, большая концентрация примесей в эмитерной области с целью увеличения σ_p сокращает время жизни неосновных носителей в этой области. Очевидно, в этом случае, как и во многих других, приходится искать какое-то компромиссное решение вопроса. Тем не менее величина γ в промышленных триодах может превышать 0,995.

21-2. ВХОДНОЕ ПОЛНОЕ СОПРОТИВЛЕНИЕ ЭМИТТЕРА

Характеристика уединенного p-n перехода в диоде согласно уравнению выпрямителя в гл. 19 есть

$$I = I_{\text{нас}}(1 - e^{eU_a/kT})\ a/см^2, \qquad (21\text{-}2)$$

где

$$J_{\text{нас}} = -\frac{kT}{e}\cdot\frac{b}{(1+b)^2}(\sigma_i^2)\left(\frac{1}{\sigma_n L_p}+\frac{1}{\sigma_p L_n}\right)\ a/см^2. \qquad (21\text{-}3)$$

Если прямое смещение имеет величину, большую 0,1 *в*, то экспоненциальный член в уравнении (21-2) становится значительно бо́льшим единицы и определяет форму характеристики в прямом направлении. Более того, если *p*-область перехода имеет гораздо бо́льшую концентрацию примесей, чем *n*-область, то второй член в скобках в уравнении (21-3) пренебрежимо мал по сравнению с первым.

Уравнение для прямого тока через переход в таком случае примет приближенно вид:

$$I_{\text{пр}} \approx \frac{kT}{e} \cdot \frac{b}{(1+b)^2} (\sigma_i^2) \frac{1}{\sigma_n L_p} e^{eU_a/kT} \; a/см^2.$$

Если же этот переход представляет собой эмиттер *p-n-p* триода с коллектором, находящимся при обратном смещении на расстоянии l, то прямой ток перехода умножается еще на величину L_p на основании положений, приведенных в предыдущем параграфе, и

$$I_{\text{пр (триода)}} \approx \frac{kT}{e} \cdot \frac{b}{(1+b)^2} \cdot \frac{\sigma_i^2}{\sigma_n l} e^{eU_э/kT} \; a/см^2.$$

Прямое полное сопротивление переменному току этого эмиттера $dU_э/dI_{\text{пр}}$ получается в таком случае путем дифференцирования, откуда

$$\frac{dU_э}{dI_{\text{пр}}} = r_э \approx \frac{kT}{e} \cdot \frac{1}{I_{\text{пр}}} \; ом/см^2. \qquad (21\text{-}4)$$

Таким образом, мы видим, что полное сопротивление эмиттера зависит от тока через эмитерный переход при прямом смещении. Интересно отметить, что это полное сопротивление эмиттера не зависит от времени жизни частиц, концентрации примесей или других величин, которые принято обычно считать параметрами прибора. Зависимость его от тока при прямом смещении делает $r_э$ скорее параметром цепи, величина которого может сильно меняться в зависимости от выбора остальных параметров цепи и, следовательно, рабочей точки по постоянному току. Входное дифференциальное сопротивление этого транзистора можно в таком случае получить, используя любое уравнение для входного полного сопротивления из гл. 11, соответствующее конкретной схеме цепи.

21-3. ЭФФЕКТИВНОСТЬ ПЕРЕНОСА ЧАСТИЦ ОТ ЭМИТТЕРА К КОЛЛЕКТОРУ

Ток неосновных носителей, идущий из базы через коллекторный переход, меньше тока этих частиц, проходящего в базовый слой через эмитерный переход; эта разница зависит от быстроты, с которой неосновные носители исчезают вследствие рекомбинации при перемещении через базу. Коэффициент переноса, обозначаемый β, определяется как отношение тока неосновных носителей, приходящих к коллектору, к току неосновных носителей, выходящих из эмиттера в базу. Коэффициент переноса можно рассчитать, использовав уравнение непрерывности для случая, показанного на рис. 21-3 кривой *2*, изображающей зависимость от расстояния концентрации неосновных носителей в базовом слое триода при прямом смещении эмиттера. Эта кривая, в первом приближении имеющая постоянный наклон в базовом слое при $l \ll L_p$, на самом деле несколько выгнута вниз, так что градиент концентрации немного меньше при абсциссе коллектора ($x=l$), чем при абсциссе эмиттера ($x=0$); эта кривизна есть следствие рекомбинации дырок на пути их перемещения. Эмитерный и коллекторный токи неосновных носителей пропорциональны этим градиентам, и по определению

$$\beta = \frac{dp/dx \text{ при } x=l}{dp/dx \text{ при } x=0}.$$

Уравнение непрерывности для этого случая (см. пример 3 из гл. 18) имеет вид:

$$0 = D_p \frac{d^2p}{dx^2} - \frac{p-p_0}{\tau_p}.$$

Его решение для нашего случая, когда x ограничено в любом направлении, есть

$$p - p_0 = Ae^{x/L_p} + Be^{-x/L_p}.$$

Дифференцируя уравнение, чтобы получить dp/dx, учитывая граничные условия, что при $x=0$ $p=p_1$, а

при $x=l$ $p=0$, и решая уравнение алгебраически, получим:

$$\beta=\frac{2(p_1-p_0)+p_0(e^{l/L_p}+e^{-l/L_p})}{(p_1-p_0)(e^{l/L_p}+e^{-l/L_p})+2p_0}.$$

Если $p_1 \gg p_0$ и $l \ll L_p$, то

$$\beta \approx \operatorname{sech}\left(\frac{l}{L_p}\right) \approx 1 - \frac{1}{2}\left(\frac{l}{L_p}\right)^2. \qquad (21\text{-}5)$$

Чтобы иметь представление о величине β у практически используемых транзисторов, можно рассмотреть более или менее типичный случай $L_p=0{,}25$ *мм* и рассчитать β: если $l=0{,}1$ *мм*; 0,05 *мм*; 0,025 *мм*; 0,01 *мм*, то $\beta=0{,}92$; 0,98; 0,995; 0,9992, соответственно.

Поскольку работа транзистора ограничивается частотами, более низкими, чем те, при которых время переноса неосновных носителей через базовый слой сравнимо с периодом сигнала, развитая здесь картина явлений переноса, очевидно, справедлива. Однако при работе на более высоких частотах следует принимать во внимание происходящие изменения величины и фазы β. Как уже предварительно обсуждалось в гл. 12, механизм ограничения максимальной частоты в транзисторах заключается в потере сигнала от рассеяния сгустков заряда, которое имеет место, когда соответствующие сигналу колебания концентрации неосновных носителей проходят через базовый слой. Аналитически эта потеря должна быть свойством коэффициента переноса β, который, как мы увидим, представляет собой одну из составных частей коэффициента передачи тока: α.

Как показано в примере 2 гл. 18, прямоугольное распределение избыточного заряда неосновных носителей, введенных в дрейфовое пространство, размазывается со временем в пятно, имеющее форму кривой ошибок Гаусса. То же происходит с группами эмитированных неосновных носителей, перемещающихся через базовый слой транзистора, независимо от того, имеют ли они форму прямоугольных или синусоидальных колебаний. Таким образом, энтропия информации сигнала увеличивается во время переноса частиц и при этом процессе теряются или относительно спадают по интенсивности высокочастотные составляющие сложной волны.

Аналитическое исследование этих процессов [Л. 1] дает для β уравнение, справедливое для высоких частот и сводящееся к уравнению (21-5) при низких частотах:

$$\beta \approx \operatorname{sech}\left[(1+j\omega\tau_p)^{1/2}\frac{l}{L_p}\right] \approx$$

$$\approx 1 - \frac{l^2}{2L_p^2} - \frac{j\omega l^2}{2D_p} \approx$$

$$\approx \beta_0 - \frac{j\omega l^2}{2D_p} \approx$$

$$\approx \frac{\beta_0}{1+\frac{j\omega l^2}{2D_p}}, \qquad (21\text{-}6)$$

где β_0 — значение β из уравнения (21-5) для низких частот. Из этого уравнения видно, что с увеличением частоты β уменьшается, достигая величины $\beta_0/\sqrt{2}$ при частоте

$$\left|\frac{j\omega l^2}{2D_p}\right| \approx 1 \text{ или } f \approx \frac{D_p}{\pi l^2}.$$

Эта частота в гл. 12 была определена как предельная частота (частота отсечки).

21-4. УСИЛЕНИЕ ТОКА В КОЛЛЕКТОРНОМ ПЕРЕХОДЕ

В плоскостных транзисторах, работающих в «обычных» условиях, коэффициент коллекторного умножения почти равен единице, т. е. на каждый неосновной носитель, проходящий через базу и прибывающий к коллекторному барьеру, во внешней цепи проходит заряд *e*. Как уже говорилось в гл. 10, это положение не выполняется для транзисторов с точечным контактом, в коллекторном переходе которых возникает вторичное явление, приводящее к усилению первичного тока обычно в 2 или 3 раза. Более того, из вида характеристик на рис. 13-5 ясно, что даже в случае плоскостных транзисторов коэффициент коллекторного усиления тока превышает единицу, если рабочая точка (коллекторное смещение) находится в верхней части области насыщения достаточно близко к напряжению пробоя.

Это усиление тока в коллекторном p-n переходе дается эмпирическим уравнением [Л. 2]

$$M = \frac{1}{1 - \left(\frac{U_к}{U_{к.п}}\right)^n}, \qquad (21\text{-}7)$$

где $U_к$ — напряжение смещения на коллекторе; $U_{к\,п}$ — напряжение пробоя коллектора, а n — константа полупроводника.

Для резких переходов, в которых изменение концентрации примеси от максимальной донорной к наибольшей акцепторной происходит на небольшом расстоянии по сравнению с толщиной слоев пространственного заряда барьера и для которых концентрация примесей гораздо больше с одной стороны, чем с другой, экспериментально были получены следующие приблизительные значения величины n:

Полупроводник	Тип перехода	Неосновные носители, за счет которых происходит усиление	Примерное значение n
Германий	p-n+	Электроны	6
»	p+-n	Дырки	3
Кремний	p-n+	Электроны	2,5
»	p+-n	Дырки	2,5

В базе транзисторов с выращенным переходом концентрация примесей обычно выше, чем в коллекторе, на целый порядок или еще больше. Так, в плоскостном транзисторе с выращенной p-n-p структурой коллекторный ток утечки при обратном смещении ($I_{к.о}$) в основном представляет собой ток электронов, идущих из коллектора в базу. С другой стороны, ток, обусловленный приходом к коллектору неосновных носителей, инъектированных эмиттером в базовый слой, представляет собой ток положительных дырок, проходящих из базы в коллектор. Именно эта составляющая тока представляет интерес для работы транзистора, и показатель n в уравнении (21-7) должен быть подобран с учетом этого обстоятельства.

В коллекторе транзисторов со сплавными переходами концентрация примесей гораздо больше, чем в базе.

В этом случае частицы, преобладающие в обратном токе утечки коллектора, те же, что и участвующие в переносе, т. е. частицы, являющиеся неосновными носителями в базе.

Поэтому германиевый транзистор *p-n-p* типа как с выращенным переходом, так и со сплавным переходом при напряжении смещения коллектора, вдвое меньшем напряжения пробоя, имеет коэффициент усиления около 112%. Подобным же образом германиевый *n-p-n* транзистор имеет коллекторный коэффициент усиления всего лишь около 101,5%. Каким бы незначительным ни казался такой коэффициент, его влияние велико. Как мы увидим дальше, усиление коллекторного тока — это один из факторов, определяющих α транзистора, а мы убедились уже в том, что при использовании транзисторов, например, в цепях с общим эмиттером полный коэффициент передачи тока из базы в коллектор в большой степени зависит от разности между величиной α и единицей. Увеличение смещения на коллекторе в такой цепи может привести к тому, что α станет больше единицы и цепь окажется нестабильной.

Поэтому, работая с плоскостными транзисторами в режиме малых сигналов, следует поддерживать напряжения на коллекторе ниже напряжения, при котором M становится заметно бóльшим единицы, а изменения M в зависимости от колебаний коллекторного смещения делаются опасными.

Физика явлений при коллекторном усилении заключается в появлении новых пар электронов и дырок в сильном поле коллекторного перехода в результате ударного возбуждения. Неосновной носитель, попадающий в коллекторный барьер, ускоряется действующим в нем полем. Полученную в этом поле энергию он может рассеять в результате столкновений с атомами решетки, теряя при каждом столкновении столько энергии, сколько он только что получил в поле. С другой стороны, он может в результате необычно длинного свободного пробега или в результате ряда последовательных упругих столкновений с атомами решетки накопить достаточную энергию для создания новой пары носителей — нового свободного электрона и новой свободной положительной дырки. Это осуществляется путем возбуждения электрона валентной зоны и переноса его в зону проводи-

мости с затратой энергии не менее eV_g. Статистическая вероятность такого производства новых пар свободных частиц возрастает по мере увеличения поля в барьере, что и приводит к такому умножению заряда и тока, которое следует из уравнения (21-7).

21-5. ВЕЛИЧИНА α И ОПРЕДЕЛЯЮЩИЕ ЕЕ МНОЖИТЕЛИ

Рассмотрим теперь α — коэффициент передачи тока и различные определяющие его множители. Описывая его словами, скажем, что изменение коллекторного тока равно первичному изменению эмитерного тока, умноженному на величину, равную доле инъектированного в базу тока неосновных носителей в общем эмитерном токе γ, умноженному на коэффициент переноса β, равный доле этих неосновных носителей, избежавших рекомбинации при движении через базу, и умноженному на коэффициент коллекторного умножения M, если он играет роль в коллекторных процессах.

Аналитически

$$|i_к| = i_э(\gamma\beta M)$$

или

$$\alpha = \left|\frac{i_к}{i_э}\right|_{U_к} = \gamma\beta M, \qquad (21\text{-}8)$$

где три множителя γ, β и M были рассмотрены нами индивидуально как для описания физики явлений, так и для установления их зависимостей от рода материала и геометрических размеров конструкции.

Очень удобно изобразить с помощью диаграммы различные приращения составляющих тока, идущего в цепи триода с *p-n-p* переходами, при небольшом положительном приращении тока $i_э$ в направлении слева направо, появляющемся в качестве сигнала на эмитерном переходе. Такая диаграмма приведена на рис. 21-4. Стрелки показывают не условное направление тока, а направление движения носителей заряда, названных рядом с каждой стрелкой.

На этой диаграмме эмитерный ток $i_э$ за эмитерным переходом в базе обнаруживается уже как сумма дырочного тока $\gamma i_э$ и электронного тока $(1-\gamma)\ i_э$. После про-

хождения через базовый слой поток дырок уменьшается на коэффициент потерь за счет рекомбинации β и дает у коллектора ток $\gamma\beta i_э$. При пересечении коллекторного барьера поток дырок снова меняется, давая в коллекторе ток $\gamma\beta M i_э$. На правом конечном контакте он превращается в электронный ток той же величины, направленный в полупроводник. На коллекторном переходе ток $\beta\gamma i_э(M-1)$, представляющий собой как бы остаток от полного тока коллектора, течет в виде электронного потока, направленного справа налево. Ток базового вывода

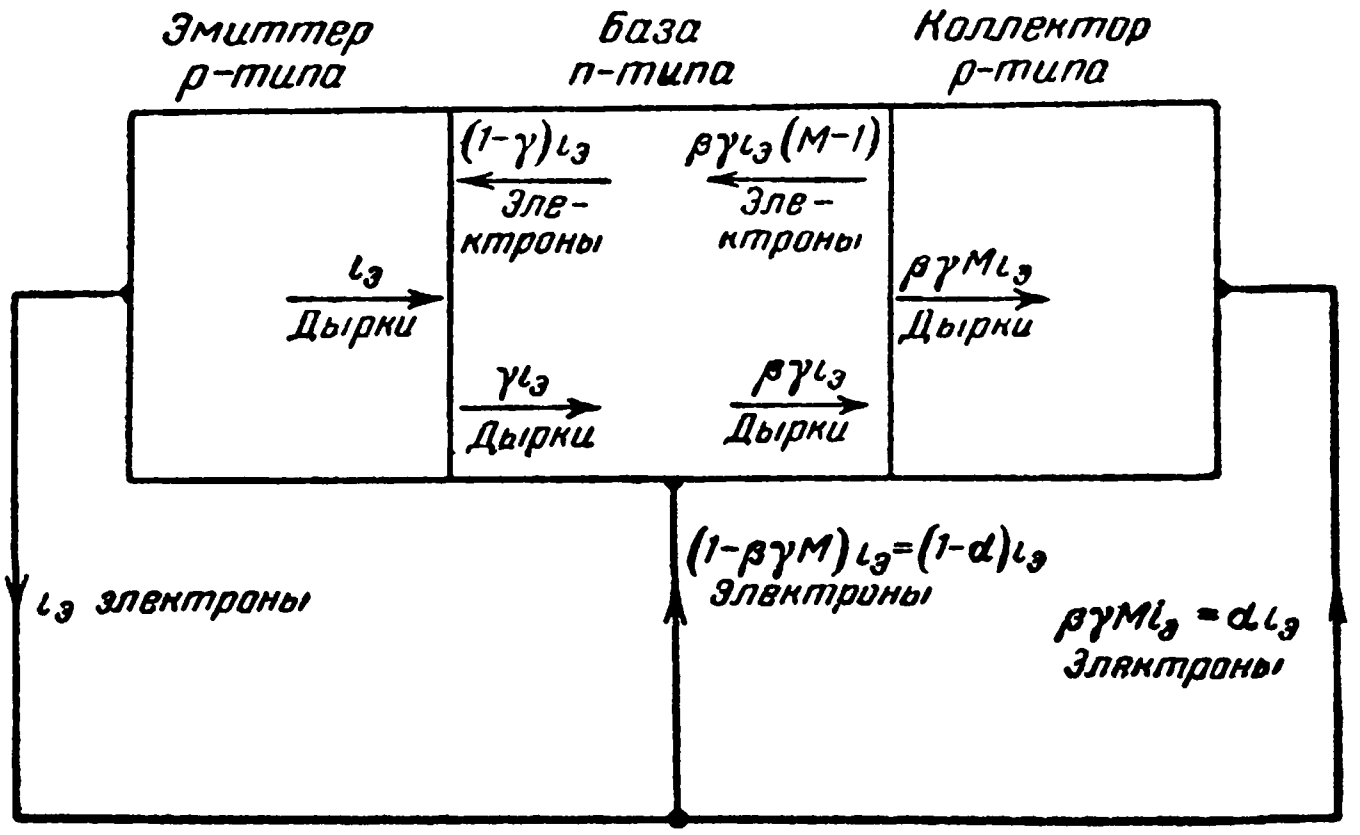


Рис. 21-4. Прирост тока в *p-n-p* транзисторе и присоединенной к нему цепи, представленный в виде электронных и дырочных составляющих.

состоит в этом случае из электронов, идущих в базовый слой для восполнения потерь из-за рекомбинации неосновных носителей и уравновешивания разностей других токов, проходящих через базу. Этот ток базового вывода $i_б$, очевидно, равен $(1-\alpha)i_э$, что необходимо для удовлетворения закону Киргофа во всех узлах и переходах в цепи. На высоких частотах различные токи подобной диаграммы должны уравновешиваться как по фазам, так и по величине.

Ясно, что частотное поведение α определяется частотным поведением отдельных множителей, от которых α зависит. Из этих множителей γ и M можно считать практически не зависящими от частоты для всех интересую-

щих нас диапазонов. В таком случае выражение для α при высокой частоте имеет вид:

$$\alpha \approx \gamma \frac{\beta_0}{1 + \frac{j\omega l^2}{2D_p}} M. \qquad (21\text{-}9)$$

Частотная характеристика транзистора в *реальной цепи* определяется не только частотной зависимостью величины α, как следует из уравнения (21-9), но также влиянием на частоту емкостей переходов. В последнем случае появляется сложная зависимость от вида цепи и включенных в нее элементов, как указывалось в гл. 12.

21-6. КОЛЛЕКТОРНЫЙ ТОК НАСЫЩЕНИЯ

До сих пор мы обращали внимание на влияние обратного тока насыщения коллектора на работу транзистора при напряжении эмиттера, большем запирающего напряжения. Этот обратный ток насыщения коллектора, обозначаемый $I_{к\,о}$, определяет остаточный ток, который проходит через транзистор (рассматриваемый как выключатель) во включенном и выключенном его положениях.

Обратный ток насыщения p-n перехода уменьшается, если на расстоянии, меньшем одной диффузионной длины, находится другой переход при обратном смещении. В случае перехода, показанного в правой части рис. 21-1, обратный ток насыщения состоит из потока *неосновных* носителей, которые возникают по обе стороны перехода вследствие тепловой генерации на расстояниях порядка диффузионных длин соответствующих неосновных носителей. С базовой стороны рассматриваемого перехода расстояние, в пределах которого могут генерироваться неосновные носители (дырки в данном случае), ограничено толщиной базового слоя l. И если l значительно меньше, чем диффузионная длина L_p, то доля коллекторного обратного тока, создаваемая дырками, возникающими в базовом слое, уменьшится в l/L_p раз. На самом же деле он уменьшится еще в 2 раза, так как эмиттер при обратном смещении отбирает половину образующихся в базе дырок. Результирующее уравнение для обратного тока насыщения коллектора в присутст-

вии эмиттера с обратным смещением отличается от уравнения для уединенного диодного перехода (19-9):

$$I_{к.о} = \frac{kT}{e} \cdot \frac{b}{(1+b)^2} (\sigma_i^2) \left(\frac{l}{2\sigma_n L_p^2} + \frac{1}{\sigma_p L_n} \right) \text{ а/см}^2. \quad (21\text{-}10)$$

ЗАДАЧА

1. Удельное сопротивление и время жизни электронов в эмиттерной области германиевого транзистора с *p-n-p* структурой составляют соответственно 0,01 *ом · см* и 20 *мксек*. В базовой области толщиной 0,005 *см* удельное сопротивление равно 1 *ом · см* и время жизни дырок 200 *мксек*. Удельное сопротивление коллекторной области 10 *ом · см*, и на коллекторном переходе лежит смещение, составляющее 25% его обратного пробивного напряжения. Рассчитайте величину α для низкой частоты и определяемую ею пороговую частоту транзистора.

ЛИТЕРАТУРА

1. Shockiey W., Sparks M., Teal G. K., *p-n* Junction Transistors, Physical Review, 1951, vol. 83, p. 151.

2. Miller S. L., Avalanche Breakdown in Germanium, Physical Review, 1955, vol. 99, p. 1234.

3. Адирович Э. И., Колотилов В. Г. и Малин Б. В., Переходные процессы в полупроводниковых триодах, «Радиотехника и электроника», 1956, 1, 8, 1052—1057.

4. Адирович Э. И. и Темко К. В., Переходная, частотная и фазовая характеристики транзистора при общем эмиттере, ЖТФ, 1957, 27, 6, 1174—1181.

5. Кузьмин В. А. и Швейкин В. И., О работе полупроводникового триода в области насыщения, «Радиотехника и электроника», 1958, 3, 10, 1269—1273.

6. Федотов Я. А., Влияние распределения примесей в базе дрейфовых триодов на их частотные свойства, «Радиотехника и электроника», 1957, 2, 10, 1261—1270.

7. Федотов Я. А., Частотные свойства плоскостных триодов, «Радиотехника и электроника», 1957, 2, 9, 1177—1199.

8. Гринберг А. А., К теории переходной характеристики полупроводникового триода, «Радиотехника», 1958, 13, 2, 51—53.

ГЛАВА ДВАДЦАТЬ ВТОРАЯ

ОПТИЧЕСКИЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ И ФОТОЭЛЕКТРИЧЕСКИЕ ЭФФЕКТЫ

Оптические свойства полупроводников не только позволили создать и практически применить фотоэлектрические приборы, но в результате надлежащим образом поставленных измерений и их понимания дали также ценные сведения о процессах внутри полупроводников и относительно расположения, плотности распределения по энергиям и заселенности энергетических уровней. Цель этой главы — рассмотреть оптические и фотоэлектрические свойства полупроводников, показать, от чего они зависят, и выяснить, какие сведения о полупроводниках можно получить с помощью оптических и фотоэлектрических измерений. Мы увидим, что фотоэлектрические свойства фотоэлементов являются естественным следствием электронных свойств полупроводника, обсуждавшихся в предыдущих главах.

22-1. О КВАНТОВОЙ ПРИРОДЕ СВЕТА

Прежде чем исследовать явления, вызванные действием света, необходимо ознакомиться с природой самого света. История науки о свете связана со множеством превратностей. Ньютон считал луч света потоком частиц, испускаемых источником света, перемещающихся прямолинейно и попадающих в глаз наблюдателя. Основываясь на этих положениях, ему удалось без труда объяснить явления отражения и преломления, но трудно было понять явление интерференции (кольца Ньютона). От корпускулярной теории света пришлось отказаться после опытов Юнга по интерференции света, выходящего из нескольких отверстий. Объяснить их казалось возможным, лишь предположив, что свет есть волновое явление. Аналитическая формулировка волновой теории была позже разработана Френелем, Фраунгофером и Корню, объяснившими во всех деталях явления интерференции света, дифракции света от углов, у краев и в отверстиях. После того как в конце прошлого века Максвелл создал электромагнитную теорию, казалось, что раз и навсегда следует считать свет электромагнитными волнами, длины которых лежат в пределах примерно одной октавы

спектра электромагнитных излучений с длинами волн между $\sim$0,4 и $\sim$0,8 *мк*.

Физики считали эту картину такой до тех пор, пока не были обнаружены нарушающие ее фотоэлектрические явления. В электровакуумном фотоэлементе при падении света на поверхности фоточувствительного катода из этой поверхности в окружающий вакуум вылетают электроны. Для вылета им необходима энергия, чтобы преодолеть потенциальный барьер на поверхности катода, совершив работу выхода. Для большинства фотокатодов эта энергия близка к 2—3 *эв* на каждый вылетающий электрон. Можно предположить, что электрон получает необходимую энергию при мгновенном захвате лучистой энергии в очень небольшом объеме, окружающем электрон. Плотность энергии внутри этого объема может флуктуировать статистически в зависимости от случайного наложения отдельных волн, испускаемых миллионами атомных излучателей в источнике света, пока в какой-то удачный момент в доступном для электрона объеме не сконцентрируется достаточно энергии, чтобы, получив ее, он смог выйти из поверхности катода, возможно даже сохранив избыток ее в виде кинетической энергии.

При анализе остаточных кинетических энергий испускаемых электронов мы встречаемся с первым явлением, которое нельзя объяснить на основании волновой теории света. Исследование этих кинетических энергий можно сделать с помощью прибора, показанного на рис. 22-1. Между фотокатодом и анодом (покрывающим внутреннюю поверхность элемента) прикладывается переменная по величине задерживающая разность потенциалов. Для любого даного значения этого задерживающего потенциала $U_{\text{торм}}$ только те электроны достигнут анода, кинетические энергии которых выше $U_{\text{торм}}+$ $+U_{\text{конт}}$, где $U_{\text{конт}}$ — контактная разность потенциалов между поверхностями анода и катода. Рассматривая анодный ток как функцию задерживающего напряжения, можно рассчитать распределение фотоэлектронов по их кинетическим энергиям.

Согласно статистической картине, которая рассмотрена в последнем параграфе этой главы, можно ожидать, что испускаемые электроны имеют широкий диапазон кинетических энергий. При более высокой интенсивности света это распределение должно было бы содержать от-

носительно большее число электронов с большей скоростью. Более того, световые волны любой длины должны создавать фотоэлектронную эмиссию, если их интенсивность достаточно высока. Однако экспериментальные данные противоречат этим предположениям. Распределение кинетических энергий на самом деле отличается от того, которое можно вывести на основе интенсивности света и рассуждения о вероятностях. В действительности

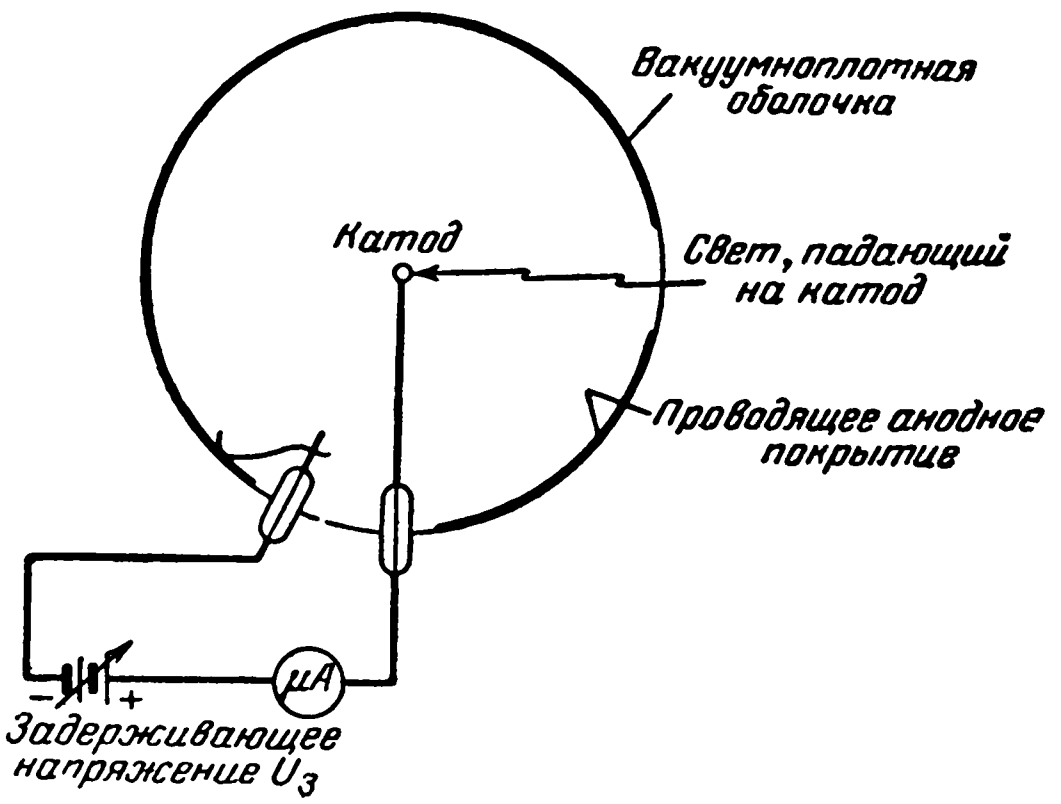


Рис. 22-1. Остаточные кинетические энергии фотоэлектронов можно измерить методом задерживающего потенциала.

это распределение вовсе не зависит от интенсивности света. Максимальная кинетическая энергия закономерно увеличивается при использовании света со все более короткой длиной волны. Единственное, что меняется в зависимости от интенсивности света, — это количество электронов, испускаемых в секунду. Кроме того, для каждого материала, из которого сделан катод, существует конкретная длина волны света, называемая пороговой длиной, при которой излучение бóльших длин волн не создает фотоэмиссии вообще независимо от того, насколько она интенсивна. Некоторые из этих положений иллюстрированы на рис. 22-2. Вторая трудность, которую волновая теория света, по-видимому, не в состоянии объяснить,—это каким образом необходимая для вырывания электрона энергия 2—3 *эв* реализуется в очень малом доступном для электрона объеме. Рассмотрим пример, когда лампа 150 *вт* находится на рас-

стоянии 1 *м* от электровакуумного фотоэлемента. Предположив, что к. п. д. лампы в излучении фотоэлектрически активного света равно 1%, и принимая, что оно распределено сферически равномерно, легко рассчитать, что плотность энергии на расстоянии 1 *м* от источника составляет около 2 *эв* на 1 *мм*³. Чтобы приобрести энергию, достаточную для вылета из поверхности катода, электрон, таким образом, должен перехватить всю сред-

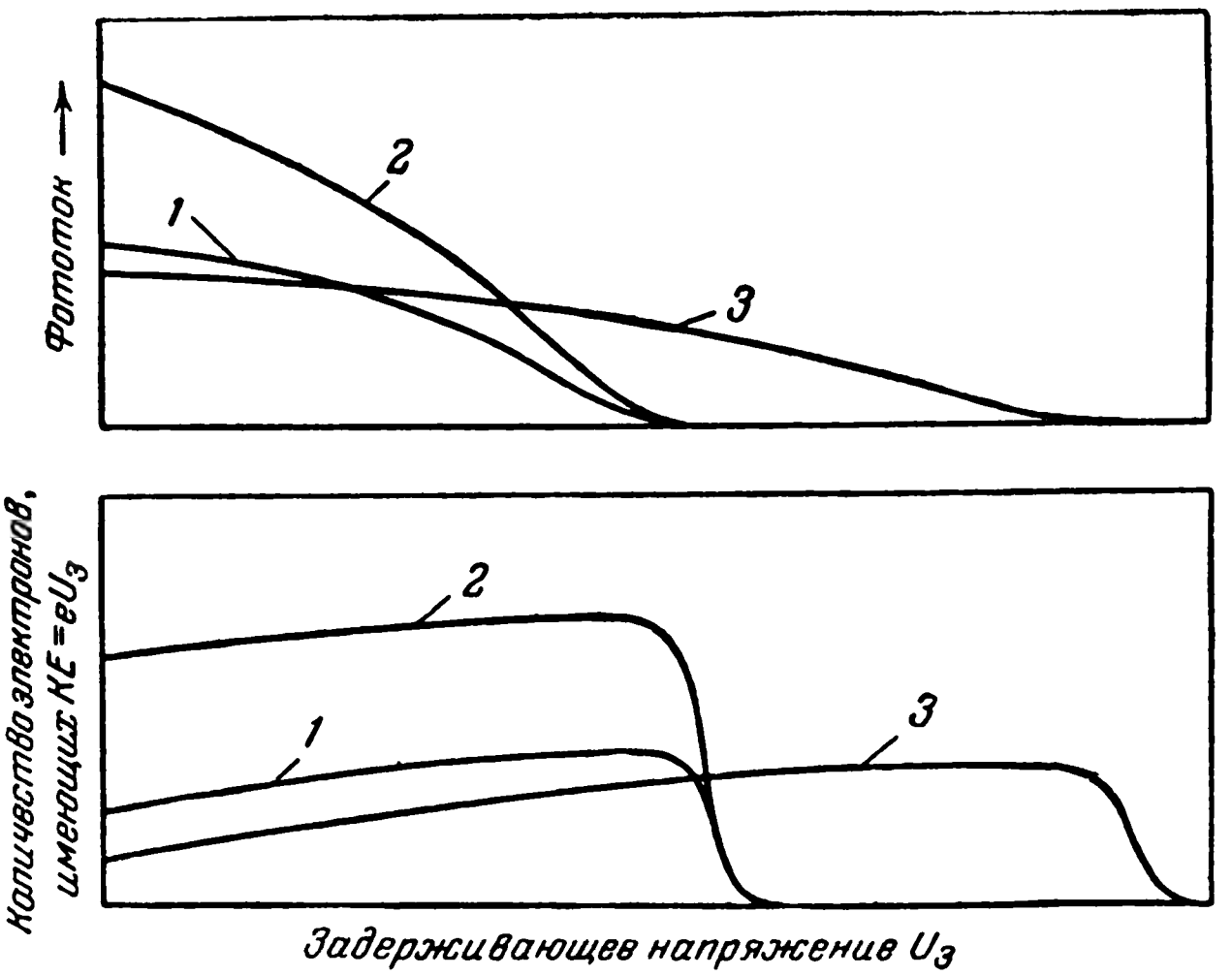


Рис. 22-2. Зависимости, получаемые при измерении методом задерживающего потенциала кинетической энергии электронов, испускаемых металлическим фотокатодом.

(*1*—определенные длина волны и интенсивность света; *2*—та же длина волны, но бóльшая интенсивность; *3*—меньшая длина волны).

нюю лучистую энергию в объеме 1 *мм*³. Даже предположив, что локальная плотность лучистой энергии очень сильно флуктуирует во времени, все же очень трудно понять, как необходимая величина энергии 2—3 *эв* может сконцентрироваться в объеме настолько малом, чтобы электрон мог поглотить ее. И тем не менее именно в этом опыте от фотокатода может быть получен фотоэлектрический ток примерно 1 *мка*/*см*².

Все эти несоответствия опытов с поведением, ожидаемым на основе простой волновой теории, легко объяснимы с помощью квантовой теории излучения. Со-

гласно последней движение электромагнитной волны в пространстве подчиняется классическим законам, за исключением того, что энергия всей волны предполагается сконцентрированной в очень малом объеме, как если бы это была частица, называемая фотоном, перемещающаяся вместе с волной. На рис. 22-3 показано плоское сечение распространяющейся в пространстве сферической волны от единичного атома, который только что испустил фотон. Процесс испускания может потребовать времени порядка 10^{-8} *сек*, а из опытов с интерферо-

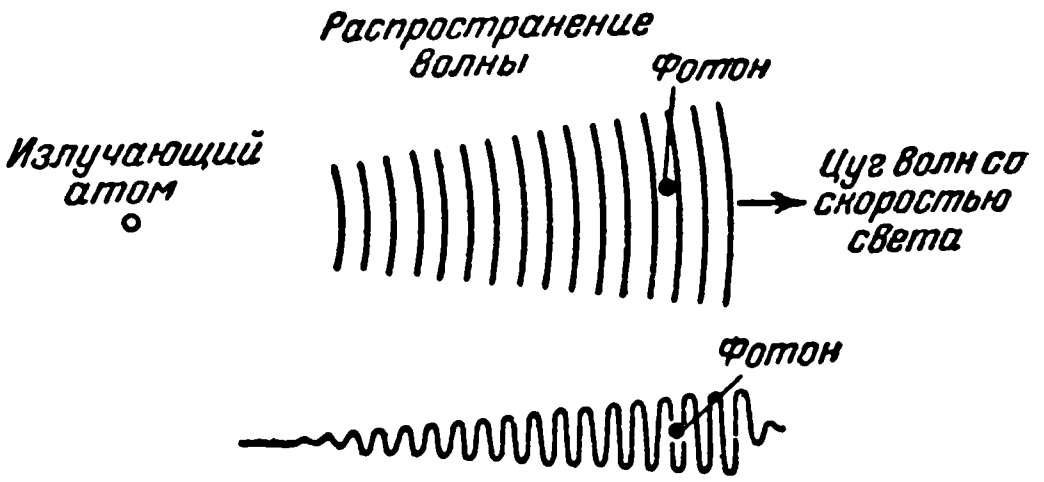


Рис. 22-3. Ожидаемый согласно волновой теории вид волны (сверху) и интенсивности поля излучения, испускаемого атомом (внизу). Энергия излучения сконцентрирована в фотоне.

метрами известно, что такая волна может иметь длину более 1 *м* и содержать несколько миллионов гребней. Таким образом, излучающий атом в момент перестройки своих электронных орбит и излучения энергии ведет себя, как дипольный осциллятор с высокой добротностью.

В нижней части рис. 22-3 показана в сечении интенсивность поля в такой волне. Вероятность нахождения фотона в определенном месте волны пропорциональна квадрату амплитуды волны в этом месте. Однако нельзя точно установить, в какой части волны фотон находится в данный момент или в каком направлении он движется от источника света, несмотря на то, что сама волна распространяется сферически. Согласно этой картине сама электромагнитная волна не обладает энергией и не переносит ее; она переносит лишь вероятность того, что где-то в этом пространстве находится фотон, несущий в крошечном объеме всю энергию, высвободившуюся в атомном процессе, породившем волну и сам фотон.

В газоразрядной трубке обычных размеров находятся миллиарды атомов, которые непрерывно возбуждаются и отдают свою энергию возбуждения в виде порций излучения. Общая картина волн вблизи такого источника очень сложна: она представляет собой наложение большого числа отдельных волн, а фотоны излучаются в виде статистически случайного потока и движутся в самых различных направлениях со скоростью света. В случае отражения, преломления, дифракции и интерференции волны ведут себя так, как предсказывают классические законы волновой механики, и фотоны движутся туда, куда направляют их волны. В этом отношении фотон и связанная с ним электромагнитная волна наделены теми же дуалистическими свойствами, что и электрон и волна Де-Бройля.

В газоразрядной трубке атомы почти все одинаковы, и излучения, которые они испускают, соответствуют электронным переходам между немногими дискретными уровнями энергии. Поэтому спектр излучения газоразрядной трубки представляет собой ряд линий с различными длинами волн, каждая из которых почти монохроматична. В случае твердого светящегося тела высокая добротность атомных излучателей нарушается непрерывными столкновениями атомов вследствие их тепловых колебаний и спектр такого твердого тела непрерывен. Однако фотонная природа излучения имеет и в этом случае тот же характер, что у газоразрядных источников, поэтому далее мы не будем делать различий между источниками света.

Волна от уединенного источника излучения может распространяться безгранично в космическом пространстве. С другой стороны, размер фотона не меняется по мере его движения с волной; следовательно, вероятность нахождения его в единице объема в волне уменьшается обратно пропорционально квадрату расстояния от источника.

Объяснение распределения кинетических энергий фотоэлектронов может быть получено на основе уравнения Эйнштейна

$$E = h\nu, \tag{22-1}$$

где E — энергия фотона; h — постоянная Планка, а ν — частота волны, связанной с фотоном.

Постоянная Планка имеет величину $6{,}62 \cdot 10^{-27}$ *эрг · сек*, или $6{,}62 \cdot 10^{-34}$ *дж · сек*. Таким образом, энергия фотона зависит только от *частоты* излучения, а не от его интенсивности. Энергия фотона называется квантом. Из уравнения (22-1) видно, что в световом поле монохроматического источника все фотоны имеют одинаковую энергию и, следовательно, должны вырывать из поверхности фотокатода фотоэлектроны с одинаковой остаточной кинетической энергией. Увеличивая интенсивность источника света, мы повышаем *число* фотонов, испускаемых в единицу времени, но не их индивидуальную энергию, а следовательно, увеличиваем ток фотоэлектронной эмиссии с поверхности катода, а не скорости испускаемых электронов. Чтобы падающий фотон был эффективным, его энергия должна быть по крайней мере равна по величине работе выхода материала катода. Вся «лишняя» энергия фотона превращается в кинетическую энергию испускаемого электрона. Это можно выразить с помощью уравнения

$$h\nu = e\varphi + KE_{\text{макс}}, \qquad (22\text{-}2)$$

где $e\varphi$ — энергия, затрачиваемая электроном при вылете из поверхности катода.

Электрон может иметь кинетическую энергию, меньшую, чем следует из этого уравнения, если он первоначально находился на энергетическом уровне, более низком, чем верхняя граница распределения электронов в частично заполненной валентной зоне металла катода. Из уравнения (22-2) видно в согласии с опытом, что по мере уменьшения частоты излучения максимальная кинетическая энергия фотоэлектронов падает. При определенной частоте (соответствующей пороговой длине волны) энергия фотона становится уже недостаточной, чтобы вырывать электроны, и фотоэлектронная эмиссия прекращается независимо от того, насколько интенсивен свет. На рис. 22-4 показаны зависимость фотоэлектронной эмиссии от частоты и характеристика максимальной кинетической энергии электронов. Сбоку представлена диаграмма энергетических уровней у поверхности металла, демонстрирующая различные величины, входящие в уравнение (22-2).

Очевидно, с помощью уравнения (22-2) можно определять работы выхода поверхностей из наблюдений по-

роговых длин волн различных материалов и проверять работы выхода, определенные из измерений термоэлектронной эмиссии в зависимости от температуры. Это предположение верно, и работы выхода для различных металлов, измеренные с помощью термоэлектронной

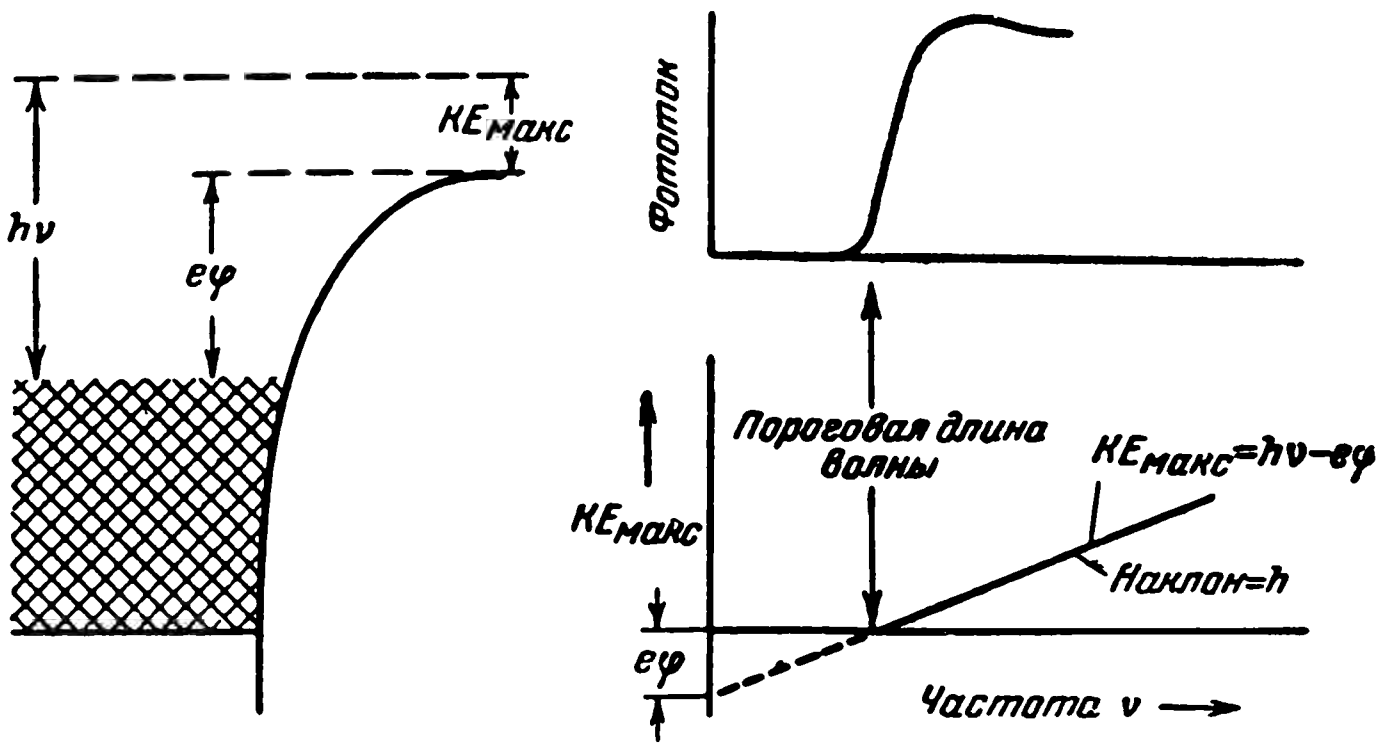


Рис 22-4 Для данной энергии фотона испускаемый фотоэлектрон имеет максимальную кинетическую энергию, если он перед выходом находился на верхнем уровне энергетического распределения электронов в катоде.

эмиссии, совпадают с величинами, измеренными методом фотоэлектрической эмиссии [Л. 1]. В тех случаях, когда это соответствие не совсем точно, причина, по-видимому, заключается в недостаточной чистоте и обработке экспериментальных поверхностей [1].

Уравнение для энергии фотона $E=h\nu$ можно представить в ином виде, выразив входящую в него энергию в электронвольтах, а длину волны — в микронах:

$$Электронвольты=\frac{1{,}23}{мк}. \qquad (22\text{-}3)$$

Таким образом, видимый зеленый свет с длиной волны приблизительно 0,55 *мк* соответствует энергии фотона примерно 2,2 *эв*. Поскольку многие поверхности фотокатодов имеют работу выхода выше, чем эта величина, соответствующие электровакуумные фотоэлементы нечув-

[1] Расхождения порядка 1—2% объясняются теоретически в связи с различием методов измерения и поэтому остаются даже при использовании самых чистых поверхностей. *Прим ред.*

ствительны к зеленому, желтому и красному свету; они чувствительны к синему и фиолетовому свету с более короткими длинами волн и фотонами более высокой энергии.

22-2. ФОТОПРОВОДИМОСТЬ

Сделав выше несколько предварительных замечаний относительно природы света, мы перейдем теперь к обсуждению фотопроводимости полупроводников. Если на поверхность полупроводника падает свет с соответствующей длиной волны, то энергия падающих фотонов затрачивается на перенос электронов из валентной зоны

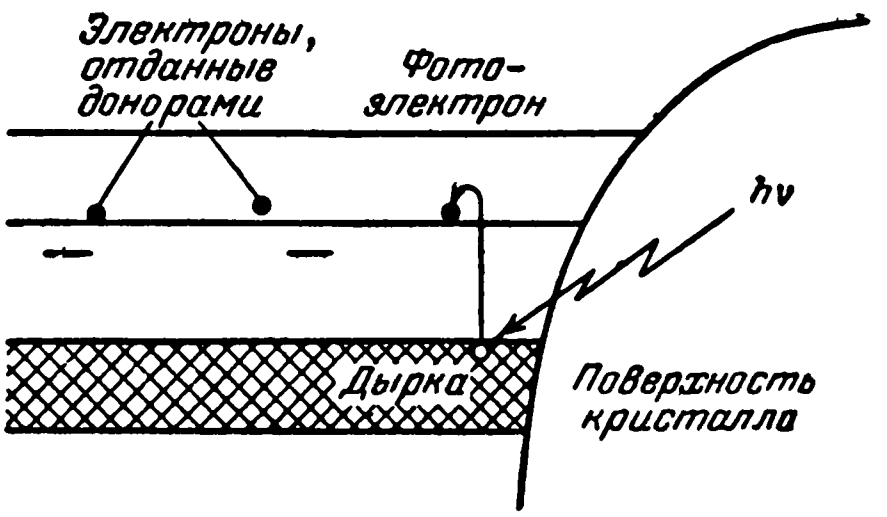


Рис 22-5 Фотопроводимость полупроводника является результатом фотоэлектрического образования новых пар электрон — дырка.

через энергетическую щель в зону проводимости. Этот процесс показан на рис. 22-5. Перенесенные таким образом электроны становятся проводящими частицами, а положительные дырки, оставшиеся в валентной зоне, вносят аналогичный вклад в возросшую проводимость облученного полупроводника.

Поскольку энергия, необходимая для возбуждения электрона и перевода его в зону проводимости полупроводника, меньше энергии, необходимой для вылета через поверхность, фотопроводимость должна вызываться световыми волнами большей длины, чем та, которая вызывает из того же материала фотоэлектронную эмиссию. Это заключение действительно верно. Фотопроводимость имеет место в полупроводнике, если кванты света имеют энергию, бóльшую, чем eV_g. Для германия величина eV_g, как было определено в результате измерений из графиков $\lg \sigma$ относительно $1/T$ в области температур,

отвечающих собственной проводимости, равна 0,7 *эв*, и отсюда пороговая длина волны для фотопроводимости должна быть примерно 1,8 *мк*, т. е. должна лежать в инфракрасной области оптического спектра. Характеристика чувствительности германиевого фототранзистора (рис. 22-6) имеет примерно такой порог, если определить

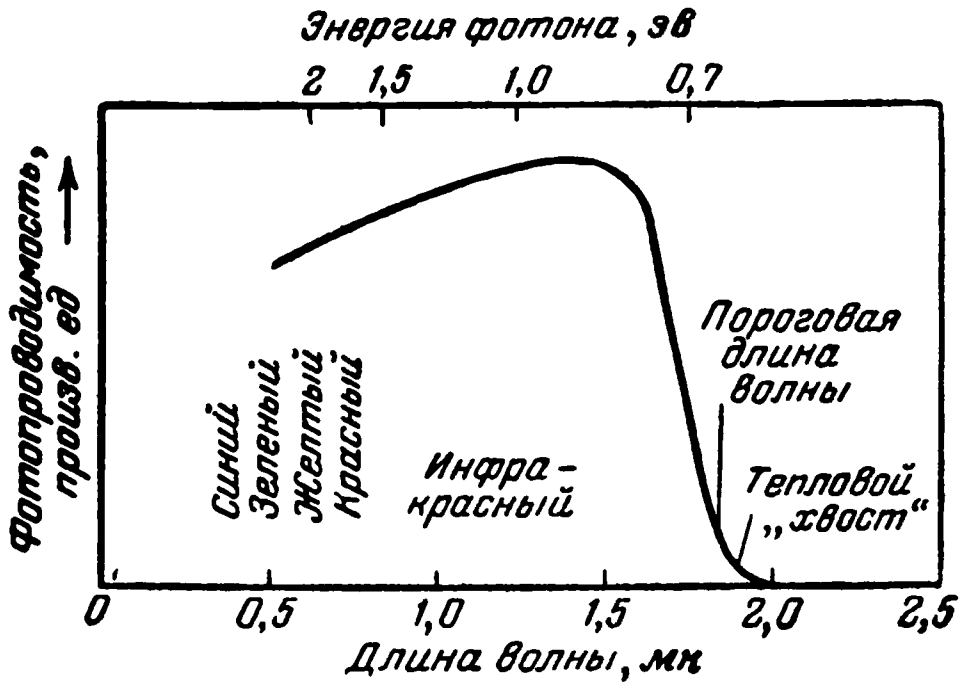


Рис 22-6. Из длинноволнового фотоэлектрического порога можно получить величину V_g полупроводника для комнатной температуры.

его, экстраполировав круто падающую кривую до пересечения с осью абсцисс.

Этот порог соответствует возбуждению электрона с самого верхнего уровня валентной зоны на самый низкий уровень зоны проводимости. Фотопроводимость при длинах волн, меньших 1,8 *мк*, соответствует энергии возбуждения электронов из низких уровней валентной зоны на более высокие уровни зоны проводимости.

На этом же рисунке видно, кроме того, что у этой кривой есть «хвост», который тянется до 1,9—2 *мк*. Два физических процесса могут объяснить небольшую фотопроводимость, обусловленную действием света, кванты которого имеют энергию, несколько меньшую величины eV_g. Во-первых, электрон может одновременно поглощать энергию фотона и энергию тепловых колебаний кристаллической решетки, получая в сумме энергию, необходимую для фотовозбуждения, хотя энергия самого фотона меньше, чем величина eV_g. Во-вторых, сама величина V_g не является абсолютно постоянной в данном полупроводнике. Тепловые колебания решетки вызывают

флуктуации мгновенной величины локальной плотности в данном объеме кристалла и приводят, следовательно, к локальным флуктуациям величины V_g, которая зависит от мгновенных расстояний между атомами кристалла. Фотон, поглощенный в области кристалла, где мгновенная величина V_g немного меньше средней величины V_g, может высвободить пару дырка — электрон, даже если его энергия несколько меньше, чем средняя величина eV_g. Хвост характеристической кривой на рис. 22-6 имеет, таким образом, тепловую природу. Он типичен для всех фотоэлектрических и термоэлектронных явлений и имеет величину порядка kT при любой температуре.

За исключением описанного выше случая и нескольких особых интересных случаев, которые описаны будут позже, можно сказать, что фотопроводимость не имеет места при длинах волн, бóльших, чем длина, соответствующая уравнению $h\nu = eV_g$. Далее, поскольку при таких длинах волн фотоны не могут тратить свою энергию на образование пар дырка — электрон, простые полупроводники сравнительно прозрачны в этих областях спектра. Однако имеет место некоторое оптическое поглощение, вследствие того что в полупроводнике есть свободные проводящие электроны и дырки [Л. 2].

Согласно классической теории [Л. 3] атмосфера из свободных заряженных частиц в среде, обладающей электрическим сопротивлением, может поглощать и преломлять электромагнитное излучение, точно так же как это происходит с радиоволнами в ионосфере. На рис. 22-7 показаны идеализированные кривые зависимости поглощения от длины волны для полупроводников с различными концентрациями примесей. За длинноволновым порогом фотопроводимости поглощение резко падает до малой величины, зависящей от толщины образца и количества свободных заряженных частиц в 1 $см^3$; при этом образцы с меньшим удельным сопротивлением поглощают сильнее, так как концентрация свободных частиц в них больше. При еще бóльших длинах волн это «классическое» поглощение снова увеличивается, как показано на рисунке. С правой стороны на том же рисунке для сравнения даны кривые, полученные экспериментально для ряда образцов кремния [Л. 4] с различной концентрацией примеси бора и, следовательно, с различ-

ной величиной удельного сопротивления получающихся полупроводников *p*-типа. На этих графиках поглощение представлено в виде коэффициента затухания, отнесенного к величине поглощения в слое образца толщиной в одну длину волны.

Из этих результатов видно, что повышение температуры полупроводника, увеличивающее концентрацию участвующих в проводимости частиц, должно повысить и поглощение в области длин волн, бóльших порога фотопроводимости. Этот эффект на самом деле имеет

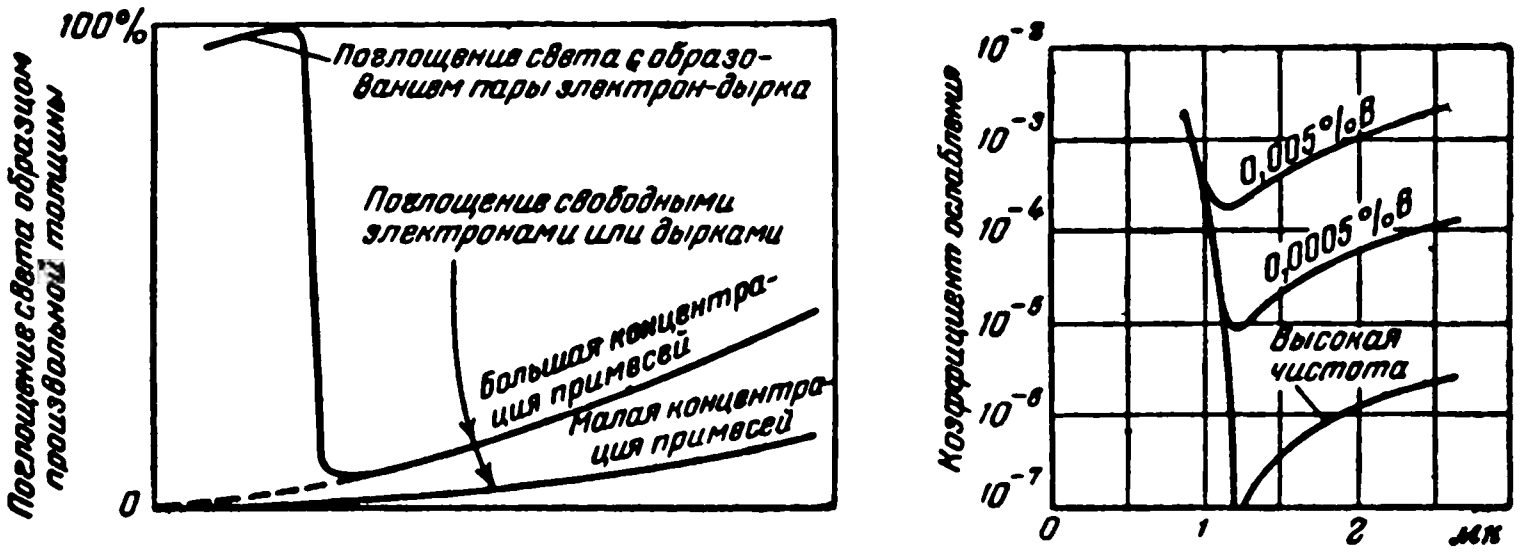


Рис 22-7. На идеализированной характеристике спектрального поглощения полупроводника (слева) видно поглощение света свободными частицами при длинах волн, бóльших пороговой Сравните с данными для кремния (справа). Коэффициент ослабления — lg величины, обратной прозрачности.

место. Закись меди имеет рубиново-красный цвет в проходящем свете, так как она имеет величину V_g, равную примерно 1,8 *эв*, и, следовательно, по уравнению (22-3) — пороговую длину волны фотопроводимости около 0,7 *мк* в оранжевой части спектра. Поглощение синего, зеленого и желтого света почти полное уже в образцах обычной толщины. Если повышать температуру закиси меди, то ее красный цвет становится более густым, затем грязновато-серым и, наконец, черным.

В фотоэлектрических процессах, при которых происходит перенос электронов из одной энергетической зоны твердого тела в другую, квантовая механика устанавливает определенные правила переходов, называемые правилами отбора. Электрон и оставляемая им дырка должны иметь в момент образования равные и противоположные по знаку импульсы [Л. 5]. Поскольку для положительной дырки энергия и импульс измеряются вниз от верхнего уровня валентной зоны, тогда как для

свободного электрона эти величины измеряются вверх от нижнего уровня зоны проводимости, закон сохранения импульсов при образовании пары электрон — дырка требует, чтобы электрон и дырка занимали уровни импульсов, симметричные относительно соответствующих границ зон. На рис. 22-8 показаны ряд таких переходов и пример «запрещенного» правилами отбора перехода. Переходы, которые по правилам отбора запрещены, тем не менее имеют место, но вероятность их тем меньше, чем сильнее нарушаются правила отбора.

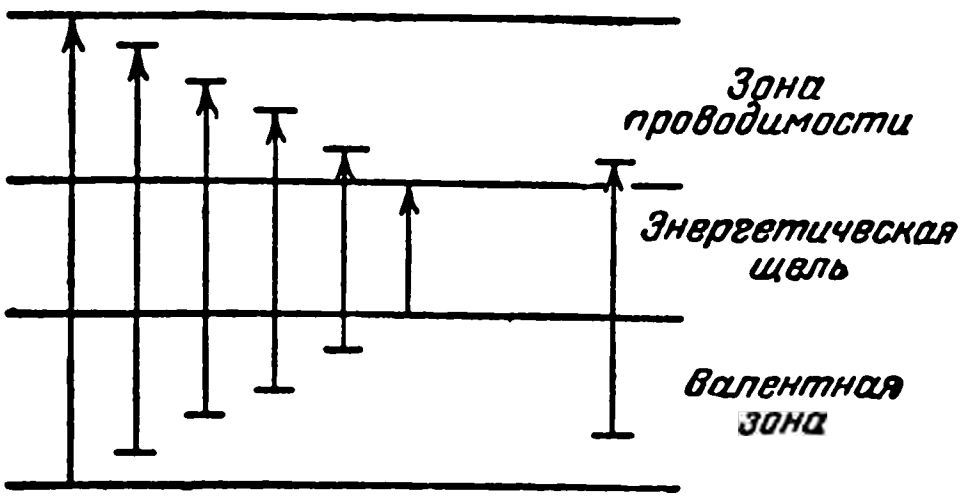


Рис. 22-8 В процессе фотоэлектрического возбуждения разрешаются лишь вполне определенные переходы.

Некоторые полупроводники сильно поглощают свет в такой области длин волн, в которой имеет место образование пар электрон — дырка, но заметной фотопроводимости при этом не наблюдается. Предполагается, что в этих случаях время жизни высвободившихся частиц очень мало. Частица, которая сразу же после освобождения рекомбинирует, не может давать такой же вклад в проводимость, как частица, остающаяся свободной долгое время.

В других полупроводниках сильное поглощение света наблюдается при длинах волн, значительно бóльших пороговой длины волны фотопроводимости. Это поглощение не сопровождается появлением фотопроводимости, но оно слишком велико, чтобы рассматривать его как классическое поглощение света свободными частицами, возникшими вследствие теплового возбуждения. Чтобы объяснить подобное явление, физики прибегли к понятию об экситоне [Л. 6] (экситон — система из взаимно связанных электрона и оставленной им дырки). В экситоне частицы не разделены полностью,

а удерживаются их собственными электростатическими полями. Таким образом, экситон — это нечто, напоминающее атом водорода, помещенный в кристалл, причем роль ядра водородного атома играет положительная дырка. Электростатическое поле этой дырки обусловливает наличие ряда дискретных уровней энергии (как в атоме водорода), из которых один или более могут находиться в энергетической щели кристалла. На рис. 22-9 показана предполагаемая для этого случая диаграмма, из которой видно, что экситон может быть

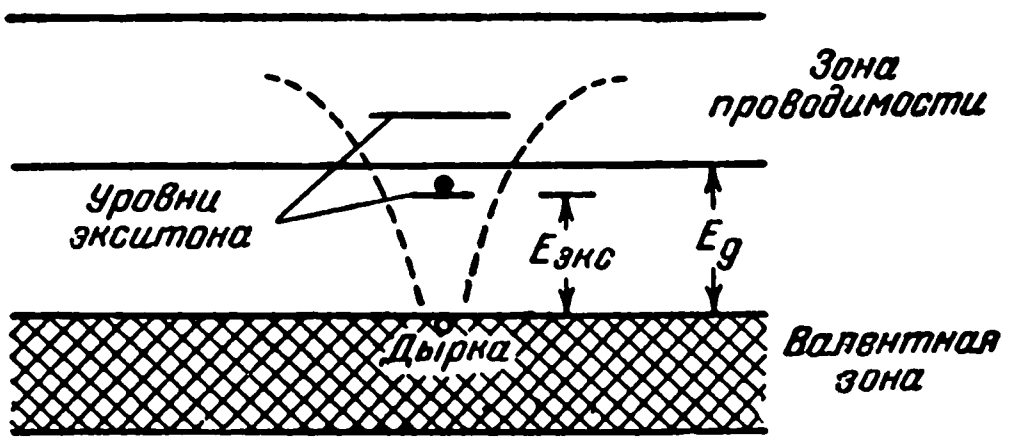


Рис. 22-9. Экситон может образоваться в результате поглощения фотона, энергия которого недостаточна для образования пары электрон — дырка.

образован фотоном с большей длиной волны, чем длина волны, необходимая для переноса электрона на нижний уровень зоны проводимости. Поскольку экситон может быть создан поглощением фотона и концентрация валентных электронов, могущих участвовать в образовании экситонов, велика, нетрудно видеть, что в соответствующем диапазоне длин волн будет иметь место сильное поглощение света. На рис. 22-10 показан типичный вид спектра поглощения твердого образца с образованием экситона в энергетической щели.

Сам экситон не является частицей, участвующей в проводимости, и не способствует фотопроводимости. На него не оказывает действие электрическое поле и его движения в кристалле и сквозь кристалл беспорядочны. Если с ним ничего не произойдет в течение короткого промежутка времени, он отдаст свою энергию решетке в виде тепловых колебаний и после рекомбинации составляющих его электрона и дырки перестанет существовать. Энергия экситона проявляется, таким образом, в нагревании решетки.

Экситонный механизм можно использовать также для объяснения случаев, в которых фотопроводимость наблюдается при длинах волн, бóльших длины, соответствующей возбуждению через энергетическую щель. В NaCl и BaO, например [Л. 7], наблюдается подобная фотопроводимость, по величине бóльшая, чем можно приписать обычному тепловому возбуждению. В этих случаях предполагают, что энергетические уровни экси-

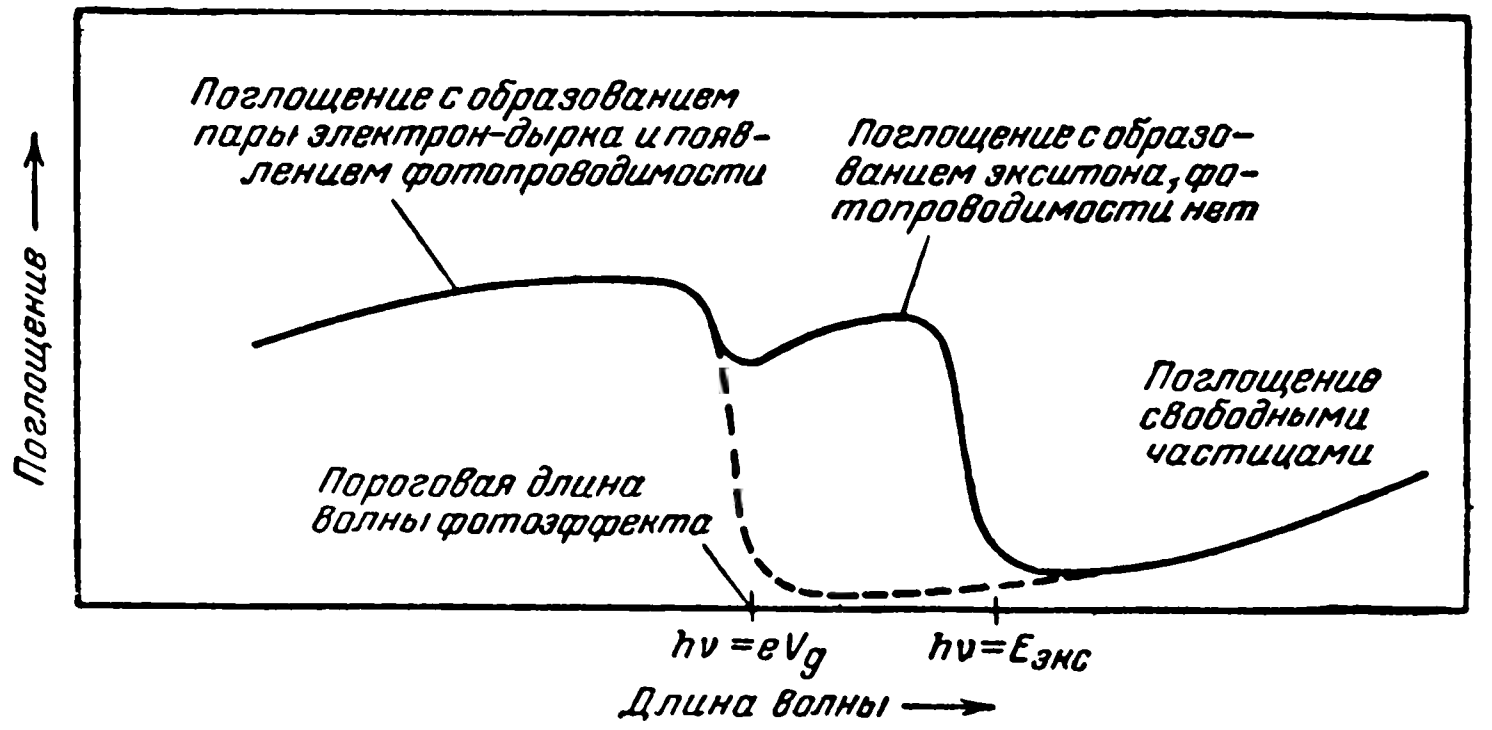


Рис 22-10. Полупроводник, энергетическая схема которого включает уровни экситона с одним значением энергии, имеет спектр поглощения, подобный показанному на рисунке.

тона лежат в энергетической щели примерно на 0,1—0,2 *эв* ниже нижней границы зоны проводимости. Поэтому после образования экситонов значительная часть из них разрушается, получая тепловую энергию кристаллической решетки, в результате чего их полная энергия становится достаточной для перехода электронов в зону проводимости кристалла. В результате этого электрон и дырка освобождаются и становятся частицами, участвующими в проводимости, несмотря на то, что первоначальная энергия фотона была меньше величины eV_g. Дальнейшее подтверждение этого мы находим в температурной зависимости фотопроводимости. С увеличением температуры увеличивается и фотопроводимость. Это можно объяснить увеличением подвижности частиц, но можно также предположить с большой вероятностью, что при более высоких температурах экситоны эффективнее отдают энергию, не успев рекомбинировать; это и

приводит к образованию большего числа подвижных носителей заряда.

В некоторых полупроводниках оптическое поглощение и фотопроводимость наблюдаются при больших длинах волн, соответствующих энергиям возбуждения донорных и акцепторных примесей. Вследствие относительно малой концентрации таких примесей по сравнению с концентрацией атомов самого полупроводника такое поглощение часто довольно трудно наблюдать. Лучше всего оно наблюдается в образцах с большой концен-

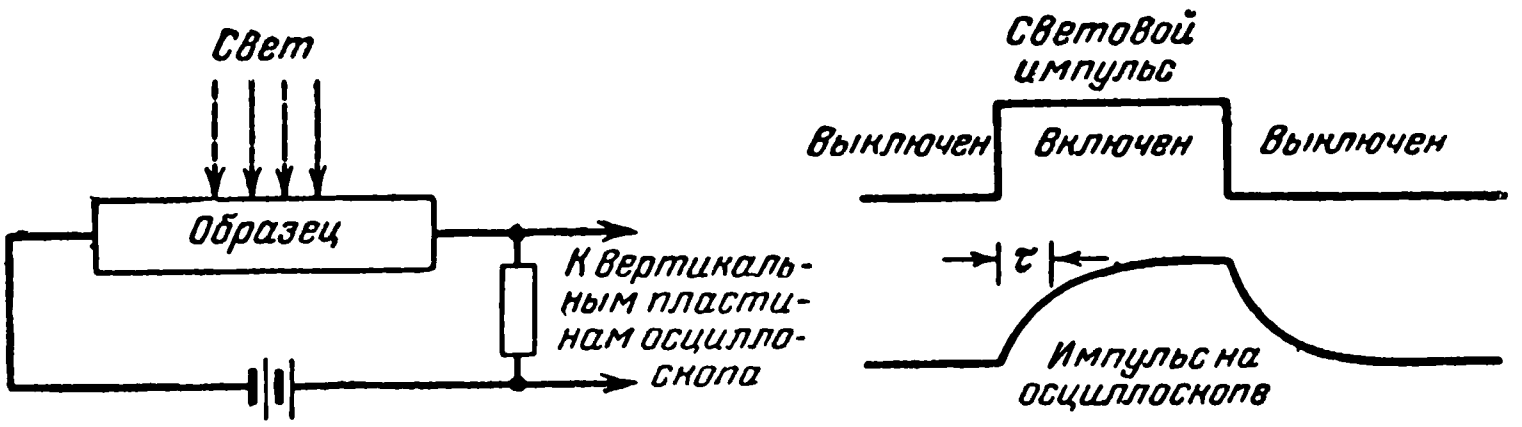


Рис. 22-11. Продолжительность жизни неосновных носителей в полупроводнике можно определить, наблюдая нарастание и спад фотопроводимости при включении и выключении света

трацией примесей и при таких низких температурах, при которых примеси термически не ионизируются. Такого рода поглощение и фотопроводимость наблюдались в кремнии [Л. 8] при длинах волн 10—20 *мк*.

Явление фотопроводимости дает возможность просто измерять продолжительность жизни неосновных носителей в полупроводниках, наблюдая быстроту возрастания фотопроводимости после включения света или ее затухания после выключения света. На рис. 22-11 показан общий вид такого экспериментального устройства. Возрастание и уменьшение величины фототока во время и после облучения образца импульсом света дают непосредственно величину τ для неосновных носителей. Применяя уравнение непрерывности, можо рассчитать, что возрастание фототока в начале импульса облучения равно

$$\Delta I = \text{const}\,(1 - e^{-t/\tau}), \qquad (22\text{-}4)$$

а спад его после конца импульса дается выражением (см. пример 1 в § 18-3)

$$\Delta I = \text{const}\; e^{-t/\tau}. \qquad (22\text{-}5)$$

Применяя этот метод, нужно пользоваться достаточно слабым светом, чтобы концентрация неосновных носителей никогда не была сравнимой с концентрацией основных носителей. Кроме того, напряжение смещения должно быть достаточно низким, а освещение сосредоточено на центральной части образца, чтобы избыточные носители не могли быть извлечены через выходные контакты. Если эти условия не соблюдаются, величина τ будет определена с ошибкой в сторону преуменьшения.

Может возникнуть вопрос относительно механизма рекомбинации. До сих пор мы допускали предположение, что рекомбинация происходит в том случае, когда свободный электрон и дырка в результате беспорядочного движения в кристалле приближаются друг к другу и электрон попадает в дырку. Этот процесс самоуничтожения должен сопровождаться испусканием фотона, уносящего разность между энергией электрона в его первоначальном состоянии в зоне проводимости и в конечном состоянии в валентной зоне.

К сожалению, несмотря на привлекательную простоту этой физической картины, только очень незначительное число рекомбинаций в полупроводниках происходит таким образом. Для того чтобы такая рекомбинация произошла, недостаточно того, чтобы электрон и дырка встретились. Должен быть выполнен закон сохранения импульса, а так как фотон практически не уносит импульса, то рекомбинация дырки и электрона может произойти только в том случае, если и электрон, и дырка имеют вначале одинаковые, но противоположно направленные импульсы. Это ограничение исключает непосредственную рекомбинацию во всех случаях, кроме очень незначительной доли столкновений электронов с дырками.

Бóльшая часть рекомбинаций в полупроводниках происходит при встрече электрона и дырки вблизи ловушки. Ловушку можно рассматривать как некое третье тело, которое может отдать или отнять импульс, необходимый для удовлетворения требованиям сохранения импульса в процессе рекомбинации. Почти все, что нарушает периодичность кристаллической решетки, может служить как такое третье тело. Примесные атомы или ионы, различные включения, незаполненные узлы в решетке, трещины, поверхности и даже различные волны

тепловых колебаний решетки могут являться «третьими телами», содействующими рекомбинации и сокращающими продолжительность жизни неосновных носителей [Л. 9]. В случае германия на 10 тыс. рекомбинаций лишь одна происходит в результате непосредственной аннигиляции с выделением фотона, уносящего всю освобождаемую при рекомбинации энергию. Подавляющая часть процессов рекомбинации происходит при участии ловушек. В процессе рекомбинации, происходящей на какой-либо ловушке, эта ловушка принимает как количество движения, необходимое для соблюдения законов сохранения импульса, так и часть энергии, освобождаемой во время рекомбинации.

Были изучены различные излучения, испускаемые электронами и дырками при рекомбинации [Л. 10, 11 и 12]. Излучение, возникающее при непосредственной аннигиляции электронов и дырок, сосредоточено в области длин волн короче пороговой для фотоэффекта. Поскольку средние тепловые кинетические энергии электронов в зоне возбуждения и дырок в валентной зоне равны $^3/_2\ kT$, прямая рекомбинация дает фотонную эмиссию, имеющую максимум интенсивности при длине волны, соответствующей энергии рекомбинации $V_g + 2(^3/_2\ kT)$. Излучение, возникающее при участии ловушек, охватывает более широкую полосу частот и распространяется до длин волн, несколько бóльших, чем пороговая длина волны фотопроводимости. Излучение рекомбинации особенно обильно у поверхности *p-n* перехода при пропускании через него прямого тока большой плотности. Пройдя переход, носители заряда становятся неосновными и рекомбинируют на расстоянии нескольких диффузионных длин от перехода.

Образец полупроводника с большой концентрацией ловушек, а следовательно, и с большой быстротой рекомбинации избыточных неосновных носителей действует так же, как область интенсивной термической генерации пар электрон — дырка. Обратный ток насыщения *p-n* перехода при обратном смещении [уравнение (19-8)] обусловлен ограниченной продолжительностью жизни неосновных носителей соседнего вещества и, следовательно, быстротой генерации новых неосновных носителей на расстоянии в несколько диффузионных длин по обе стороны от перехода. По этой причине при разра-

ботке и производстве плоскостных полупроводниковых приборов с низким обратным током насыщения необходимо принимать предосторожности, чтобы области с большой быстротой рекомбинации находились на расстоянии нескольких диффузионных длин от переходов с обратным смещением. Поэтому обработанные пескоструйным способом поверхности с припаянными контактами применяются только на концах относительно длинных стержневых структур.

22-3. ФОТОГАЛЬВАНИЧЕСКИЕ ЭФФЕКТЫ

В таком фотогальваническом элементе, как солнечный элемент, возникает э. д. с., когда свет поглощается в выпрямляющем контакте или в *p-n* переходе. На

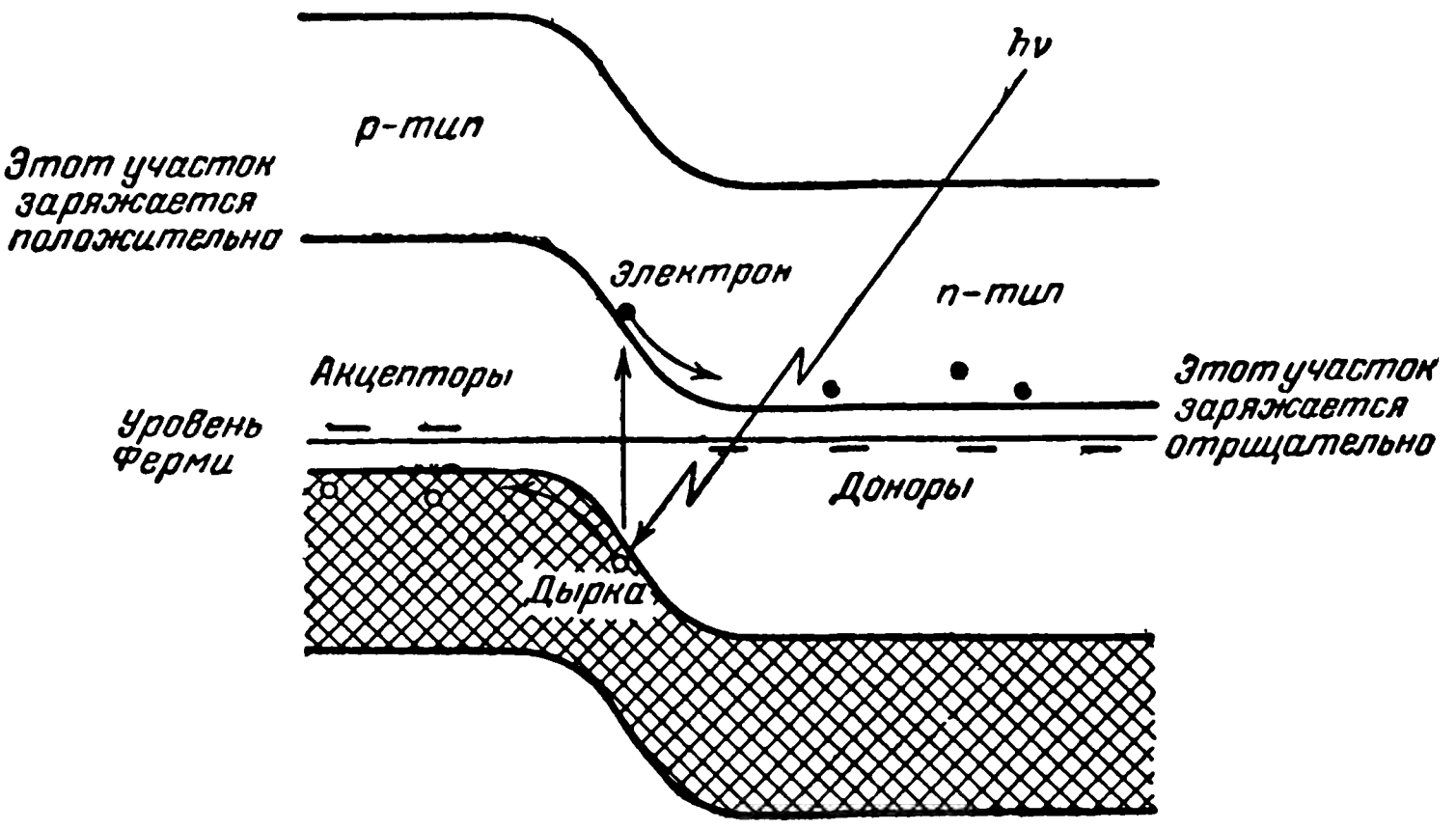


Рис. 22-12 Фото-э д. с появляется при образовании пары электрон — дырка на освещенном переходе, а затем пары разделяются полем перехода

рис. 22-12 показана схема физических явлений, лежащих в основе фотогальванического процесса в фотоэлементе с *p-n* переходом. Предполагается, что вначале переход находился в равновесии без приложенного внешнего смещения. Процесс вызывается поглощением фотона в запирающем слое и образованием пары электрон — дырка. Эти частицы разделяются внутренним полем запирающего слоя, причем электрон движется в полупроводник *n*-типа, а дырка — в полупроводник *p*-типа. Таким обра-

зом, обе области становятся противоположно заряженными и на контактах, присоединенных к *p-n* переходу, возникает э. д. с. В результате, пока переход освещен, во внешней цепи может протекать ток.

Разделение зарядов при фотогальваническом процессе ведет к заряжанию области *n*-полупроводника отрицательно, а области *p*-полупроводника положительно. Эта полярность такова, что смещение на *p-n* переходе имеет прямое направление. Если во внешней цепи нет тока, то напряжение на зажимах такой разомкнутой цепи увеличивается до тех пор, пока прямое напряжение на переходе не станет достаточным, чтобы через переход шел прямой ток, точно соответствующий быстроте освобождения новых носителей при поглощении фотонов и разделении их запирающим полем. С другой стороны, если внешняя цепь замкнута на конечную нагрузку, фототок разделяется между внешней цепью и внутренней шунтирующей барьерной цепью. Действие цепи такого вида обсуждено в гл. 8.

Для получения фотогальванического эффекта необязательно, чтобы фотоны поглощались внутри самого перехода. Фото-э. д. с. может создаваться и при поглощении света на таком расстоянии от этого барьера, что получающиеся неосновные носители диффундируют в поле запирающего слоя и благодаря этому накапливаются в соответствующих участках полупроводника. Для того чтобы такой элемент работал эффективно, необходимо, чтобы свет падал на полупроводник в пределах диффузионной длины от перехода.

Спектральная чувствительность фотогальванического элемента, конечно, такая же, как у фотосопротивления, если они сделаны из одинакового материала, так как кванты, требуемые для фотоосвобождения пар электрон — дырка, одинаковы в обоих случаях.

Во внешнюю нагрузку может быть передана большая переменная мощность при заданной амплитуде световых колебаний, если фотогальванический элемент с *p-n* переходом включен в цепь при постоянном обратном смещении. Тогда этот фотоэлемент становится фотосопротивлением с *p-n* переходом. В этих условиях внутреннее полное сопротивление перехода фотоэлемента не мало, как это характерно для прямого смещения, и единственные внутренние потери получаются в высоком со-

противлении перехода с обратным смещением, включенного параллельно нагрузке. Вследствие этого можно использовать большие сопротивления нагрузки. Те же физические процессы, которые имеют место в фотогальванических элементах с *p-n* переходом, имеют место и в подобных фотоэлементах с металл-полупроводниковым контактом.

22-4. ФОТОЭЛЕКТРОННЫЕ УМНОЖИТЕЛИ

Давно было известно, что фотосопротивления, изготовленные нанесением тонких слоев различных полупроводников на изолирующие основания, дают квантовый выход, больший единицы [Л. 3].

Для таких приборов использовались сульфид свинца, кремний, сульфид таллия и селенид кадмия. Если один падающий фотон вызывает перемещение нескольких электронов во внешней цепи, то процесс проходит по меньшей мере две стадии: 1) первичный процесс, в течение которого поглощается фотон и создается одна пара электрон — дырка; 2) вторичный процесс, при котором освобожденные первичные частицы вызывают поток других частиц. Эксперименты [Л. 14] показали, что существует два различных процесса — первичный и вторичный, но физическая сторона последнего процесса оставалась неясной до тех пор, пока не появились фотоэлектронные умножители с *n-p-n* структурой. С их помощью удалось представить себе механизм, пригодный для всех полупроводниковых фотоэлементов, имеющих общий квантовый выход, больший единицы.

Энергетическая диаграмма *n-p-n* структуры показана на рис. 22-13 для случая наложенного постоянного смещения. При этом смещение на одном из переходов обратное, в то время как на другом оно небольшое прямое. Ток между контактами внутри полупроводника создается в основном за счет переноса электронов из левой *n*-области, где они являются основными носителями, через левый запирающий слой с прямым смещением, а оттуда через правый запирающий слой с обратным смещением в правую *n*-область. Основным препятствием на пути этого потока является левый запирающий слой.

Когда освещается *p*-область, внутри нее освобождаются пары электрон — дырка. Электроны быстро диффундируют сквозь один из запирающих слоев и исчезают

из этой области. Однако положительно заряженные дырки, захваченные в валентной зоне *p*-слоя, создают в этом слое значительный пространственный заряд. Это явление аналогично накоплению пзуырьков воздуха в опрокинутом под водой стакане. Образовавшийся пространственный заряд вызывает дополнительное прямое смещение на левом запирающем слое, усиливая поток

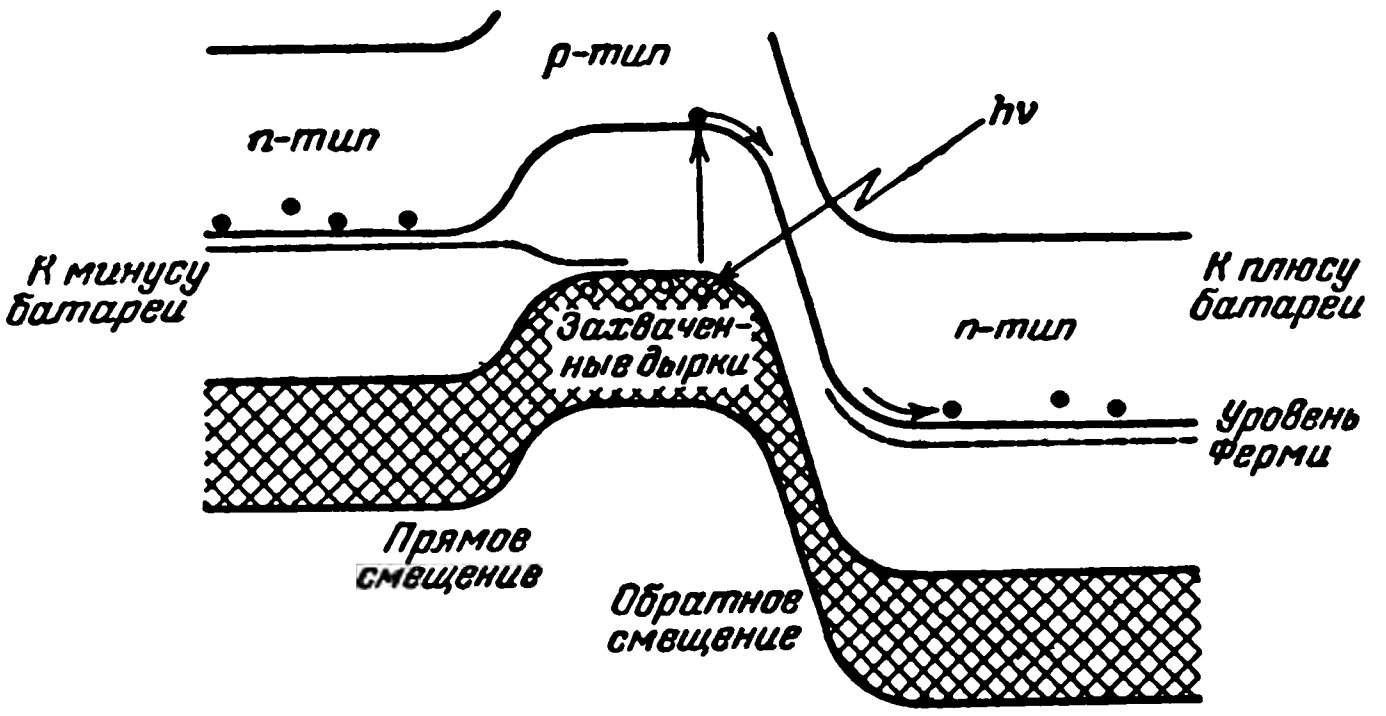


Рис. 22-13 В освещенном *n-p-n* фотоэлектронном умножителе дырки, захваченные в область «базы», понижают максимум потенциальной энергии в ней, вследствие чего большее число электронов может течь через «базу» в зону проводимости.

электронов слева направо в зоне проводимости. Бóльшая часть электронов этого потока диффундирует через *p*-область и через правый запирающий слой в правую контактную область. Некоторые электроны рекомбинируют с излишком дырок в *p*-слое, но в установившемся состоянии эта потеря компенсируется образованием новых дырок вследствие фотогенерации все время, пока действует свет.

Из рисунка видно, что освобожденный фотоэлектрически первичный заряд сам вызывает увеличение тока. Он действует как триггер или управляющее приспособление, вызывающее поток большего числа других частиц. Аналогичный случай можно наблюдать в обычном вакуумном триоде, в котором небольшой заряд управляющей сетки вызывает поток во много раз большего заряда между катодом и анодом.

То же объяснение, которое разработано для усилительного действия фотоэлемента с *n-p-n* структурой, очевидно, пригодно и для фотосопротивлений в виде тон-

ких пленок сульфида, оксида и селенида. При изготовлении таких элементов полупроводниковые слои подвергаются различным активирующим обработкам, состоящим в нагревании в различных атмосферах. Действие такой обработки заключается, по-видимому, в преобразовании некоторых областей этого слоя таким образом, чтобы материал имел свойства, отличные от свойств исходного вещества. Так, например, при частичном превращении слоя сульфида свинца в оксид свинца можно

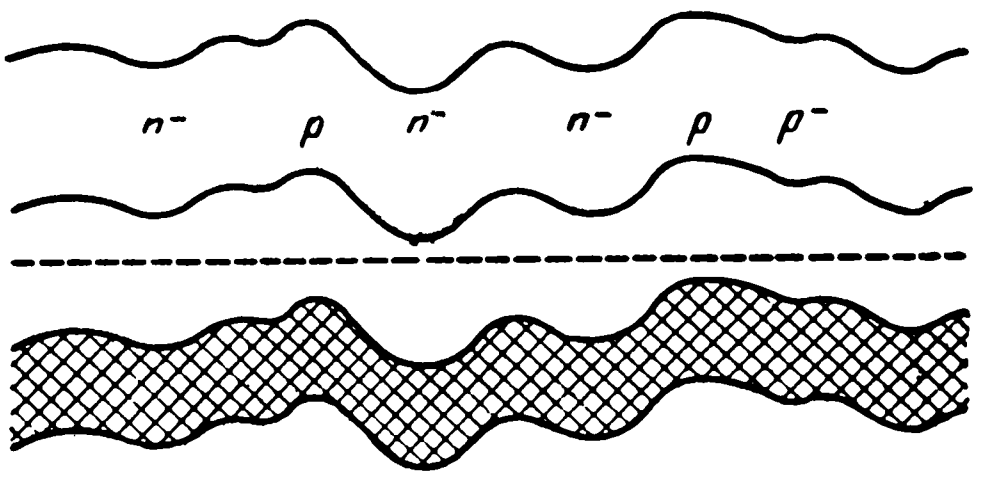


Рис. 22-14. Считается, что в фотоэлементе из сульфида свинца слой полупроводника представляет собой мозаику с хаотически распределенными умножающими ловушечными переходами.

получить слой со случайным расположением оксидных и сульфидных областей в виде мозаики. Или в слое полупроводника, который благодаря компенсации донорной и акцепторной примесей становится почти собственным, случайные флуктуации в концентрациях доноров и акцепторов в небольших участках слоя могут вызвать местные активирования n-типа или p-типа [Л. 15]. Считается, что каков бы ни был механизм образования, полупроводниковые соли этих фотоэлементов содержат беспорядочно распределенные материалы n- и p-типа или различные степени активирования состояний n- или p-типа. На энергетической диаграмме вдоль такого слоя видны расположенные в случайном порядке невысокие p-n барьерчиками или небольшие возвышения и провалы между участками типов n, n^+, p, p^+ и т. д. Такая диаграмма показана на рис. 22-14. Продольное сопротивление такого слоя определяется в основном барьерами на пути тока. Образование пар дырка — электрон при освещении слоя приведет к захвату дырок областями слоя преобладающего p-типа и захвату электронов областями

преобладающего *n*-типа. Создающиеся объемные заряды стремятся выровнять возвышения и провалы, облегчая электронам проводимости и дыркам проход через эту мозаику и уменьшая продольное сопротивление слоя

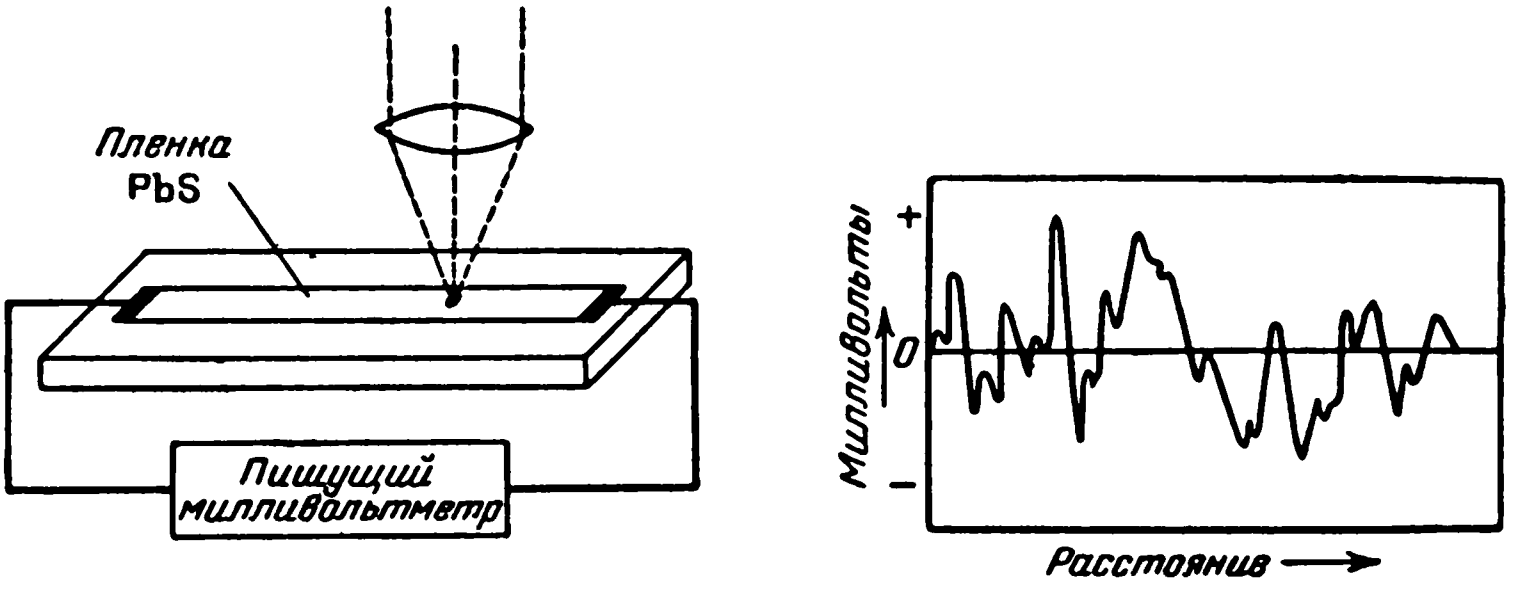


Рис 22-15. Сульфидно-свинцовая пленка, по которой перемещается маленькое пятно света, дает на выходе переменную фото-э. д. с, обнаруживая присутствие множества ловушечных переходов.

[Л. 16]. Существование границ типа *p-n*, *p-p*+ или *n-n*+ в таких пленках было показано [Л. 17] путем перемещения от одного конца к другому узкого пучка света и наблюдения возникающей на концах фото-э. д. с. Эта э. д. с. быстро меняется по знаку и величине по мере передвижения светящейся точки (рис. 22-15).

22-5. ФЛУОРЕСЦЕНЦИЯ И ФОСФОРЕСЦЕНЦИЯ В ПОЛУПРОВОДНИКАХ

Флуоренсценция и фосфоресценция — свойства материалов, присущие отнюдь не только полупроводникам [Л. 18]; они наблюдаются в твердых изоляторах, жидкостях и газах. Флуоресценцией называется свойство материала испускать свет под действием света, рентгеновых лучей, бомбардировки частицами или трения. Фосфоресценцией называется разновидность флуоресценции, задержанной некоторым внутренним механизмом, приводящим к накоплению энергии возбуждения и медленному выделению ее через некоторое время, длительность которого у разных фосфо́ров измеряется от единиц микросекунд до нескольких часов. Таким образом, флуоресценция имеет место в периоды облучения, а фосфор продолжает излучать свет и после того, как источник возбуждения удален. Флуоресценцию и фосфоресценцию

твердых тел объединяют общим названием люминесценция.

Для понимания механизма флуоресценции не требуется глубокого изучения твердых тел. В большинстве случаев она объясняется просто увеличением энергии атомов твердого тела или примесей в твердом теле вследствие поглощения излучения с переходом в более высокие состояния и немедленным (приблизительно за 10^{-8} *сек*) возвращением этих атомов к первоначальным состояниям с низкой энергией в результате излучения света. В этом отношении их поведение очень сходно с поведением газов при флуоресценции. Основное отличие заключается в том, что включение флуоресцирующих атомов в твердые вещества приводит к изменению длин волн возбуждающего и излучаемого света и расширяет линии поглощения и излучения, в результате чего образуются спектральные полосы, имеющие значительную ширину.

Обсужденное в предыдущем параграфе излучение рекомбинации считается скорее видом фосфоресценции, а не флуоресценции, поскольку возбуждение накапливается в течение некоторого времени в форме свободных электронов в зоне проводимости кристалла. Эта фосфоресценция спадает экспоненциально со временем, причем постоянная времени равна времени жизни неосновных носителей в данном кристалле.

С другой стороны, для того чтобы правильно объяснить механизм длительного послесвечения фосфо́ров, требуется объяснить накопление энергии в течение времени более долгого, чем вычисляемое, из простых соображений времени жизни. Полагают, что после того как электроны в таких фосфо́рах переводятся в зону проводимости в результате поглощения энергии возбуждения, они попадают в ловушки (захватывающие центры), где остаются в течение некоторого времени, пока не будут снова переведены вследствие повторного возбуждения в зону проводимости.

На рис. 22-16 изображена предполагаемая энергетическая диаграмма для фосфора типа сульфида цинка, а также, возможно, и для фосфора типа виллемита. На диаграмме показаны не только энергетические зоны кристалла, но и локальные уровни атомов примеси двух типов в энергетической щели. Первая из этих примесей

является захватывающим центром — ловушкой. Ловушка представляет собой примесь акцепторного характера с обычно незанятым энергетическим уровнем, расположенным несколько ниже дна зоны проводимости. Свободный, перемещающийся в кристалле электрон зоны проводимости может случайно встретить такую ловушку и быть ею захвачен вследствие понижения его энергии до локального энергетического уровня ловушки. Если этот энергетический уровень ниже дна зоны прово-

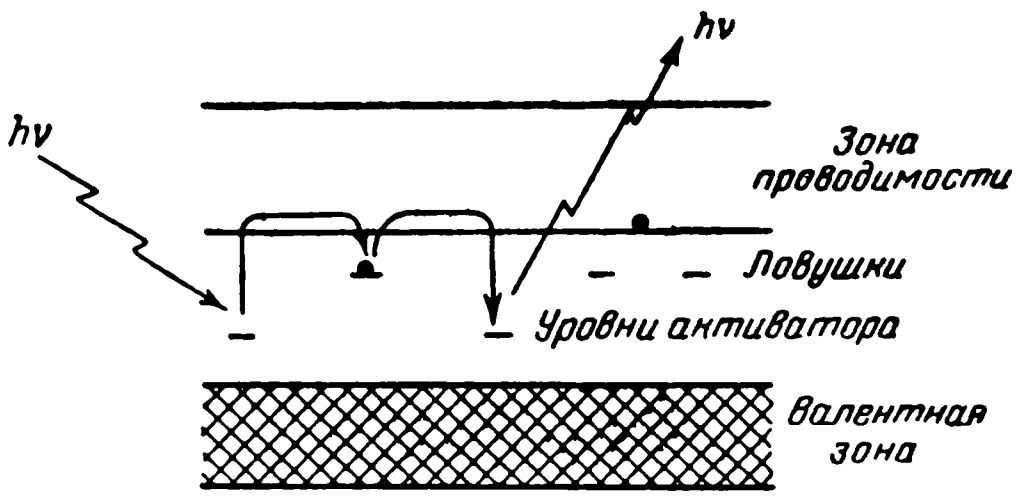


Рис. 22-16. В фосфоре из сульфида цинка электрон, перешедший с уровня активатора в зону проводимости полупроводника, попадает в ловушку, из которой он освобождается спустя некоторое время термически и снова захватывается активатором, излучая фотон.

димости более чем на несколько kT, то потребуется некоторое время, прежде чем этот электрон будет снова термически возбужден и перейдет опять в верхнюю зону. Затем электрон может рекомбинировать с положительной дыркой, испустив фотон, или же может быть снова захвачен ловушкой. По-видимому, существуют правила отбора, запрещающие электрону рекомбинировать с положительной дыркой, пока он находится на улавливающем уровне ловушки. Время пребывания электрона в захваченном состоянии определяет в основном запаздывание высвечивания фосфо́ров с длительным послесвечением. Фосфор, у которого энергетические уровни ловушек лежат значительно ниже дна зоны проводимости, имеет более длительное послесвечение, чем фосфор, у которого уровни ловушек лежат ближе к этой зоне.

На рис. 22-16 показан также тип примеси, называемый активатором. Активатором является атом донор-

ной примеси, энергетический уровень которого расположен значительно ниже дна зоны проводимости, который не может быть обычным тепловым донором. Присутствие в кристалле примеси активаторов в концентрации порядка 1/10 000 сильно увеличивает фосфоресценцию. При возбуждении фосфора некоторые электроны активаторов переходят в зону проводимости вместе с электронами из валентной зоны. Это явление, по-видимому, подтверждается тем, что спектр поглощения фосфора, содержащего активаторы, лежит в области более длинных волн, чем основной порог поглощения для чистого кристалла.

После прекращения действия источника возбуждения электронные ловушки медленно отдают захваченные электроны в зону проводимости; многие из этих электронов, если они не захватываются вновь, рекомбинируют с активаторами, а не с положительными дырками в валентной зоне. Спектр эмиссии такого фосфора, следовательно, зависит от природы активаторов и размещения их энергетических уровней в энергетической щели, а также от природы самого кристалла. Установив природу активатора, можно определить цвет излучения фосфора.

Обсуждаемый механизм предложен для объяснения поведения фосфоров типа сульфида цинка, исследованных более детально, чем другие фосфоры, и, видимо, наиболее полно понятых. Другие механизмы явлений в фосфорах были предложены для объяснения особенностей поведения других типов фосфоров, и, видимо, невозможно описать поведение всех видов фосфора с помощью одного механизма.

22-6. ФОТОЭЛЕКТРОННАЯ ЭМИССИЯ ПОЛУПРОВОДНИКОВ

Электровакуумные фотоэлементы с фотоэлектронной эмиссией, в которых полупроводники служат катодом, еще не вошли в промышленное употребление[1]. Фотоэлектронная эмиссия полупроводников изучалась, однако, как физический метод исследования с целью получения сведений по расположению и заполнению энергетических зон. Многие сведения были получены из опытов

---

[1] Широко используемые уже около четверти века сурьмяно-цезиевые и кислородно-цезиевые фотоэлементы имеют катоды из материалов, обладающих свойствами полупроводников. *Прим. ред.*

с задерживающим потенциалом. Приведем описание такого эксперимента и обсудим, что с его помощью может быть изучено, сначала для случая двух металлов, а затем для случая металла и исследуемого полупроводника [Л. 19].

Пусть из этих двух металлов изготовлен сферический конденсатор, соединенный с источником изменяемого низкого напряжения, как показано на рис. 22-1. Цен-

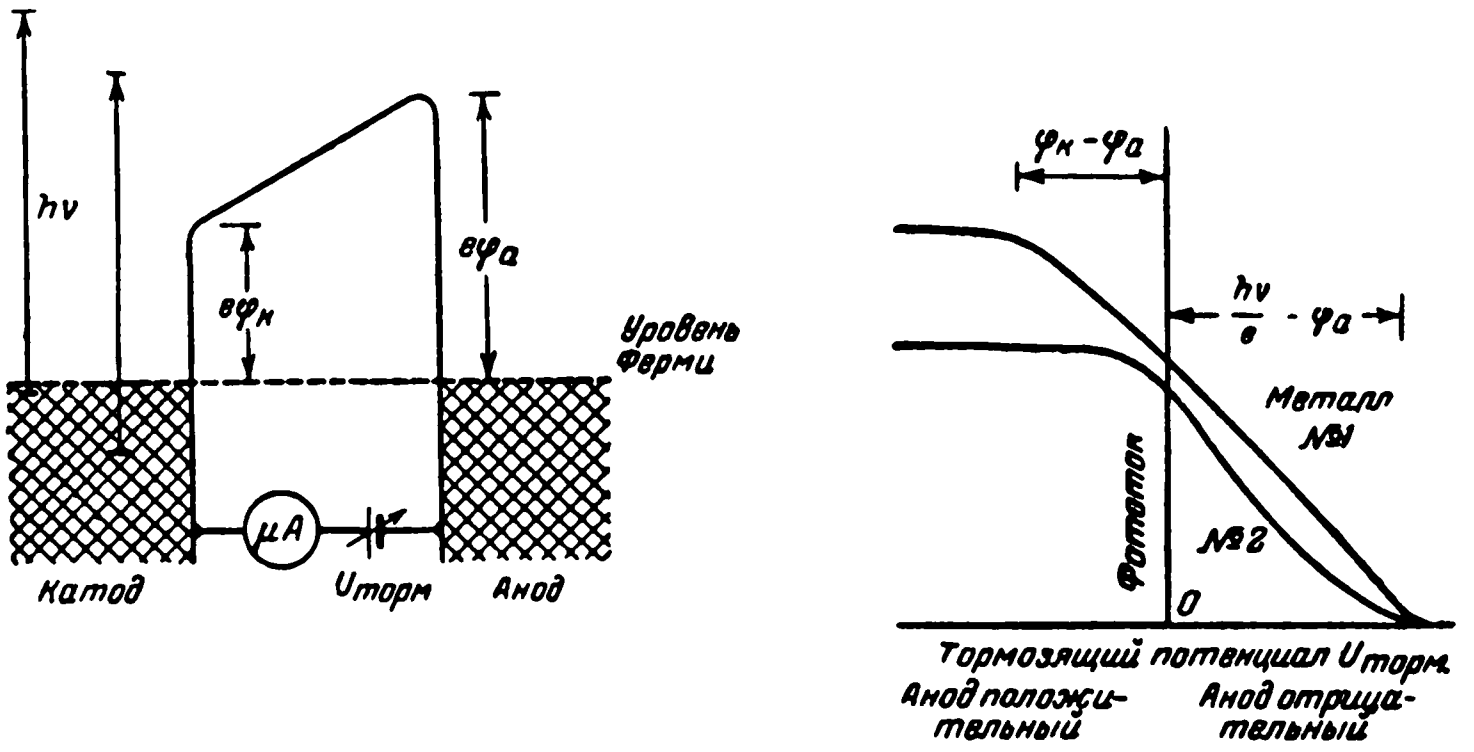


Рис. 22-17. Слева показана энергетическая схема для металлических катода и анода в экспериментальной установке, изображенной на рис. 22-1, для нулевого напряжения батареи и освещения катода монохроматическим светом. Справа — график зависимости фототока от напряжения батареи для двух катодов из различных металлов. Отсечка наступает при одинаковом $U_{торм}$.

тральный электрод, сделанный из изучаемого металла, является фотокатодом; он установлен так, что на него может падать свет. Другой металл, покрывающий изнутри бóльшую часть вакуумной оболочки, является анодным коллектором для испускаемых электронов. Пусть металл анода имеет более высокую работу выхода, чем металл катода. На рис. 22-17 приведена энергетическая диаграмма для двух таких металлов в случае отсутствия запаздывающего потенциала. Уровень Ферми на диаграмме горизонтален. Контактная разность потенциалов между этими металлическими поверхностями $\varphi_а - \varphi_к$ создает между ними внутреннее тормозящее поле.

Пусть фотокатод освещается постоянным монохроматическим светом и энергия фотонов достаточна для вы-

рывания электронов, способных преодолеть контактную разность потенциалов и достичь анода за счет избытка кинетической энергии. Пусть далее напряжение батареи увеличивается также в направлении повышения тормозящего потенциала. При этом все меньшее число электронов сохраняет способность достичь анода и фотоэлектронный ток в конце концов прекратится при тормозящем напряжении $(h\nu/e) - \varphi_a$ (если не учитывать «хвост»,

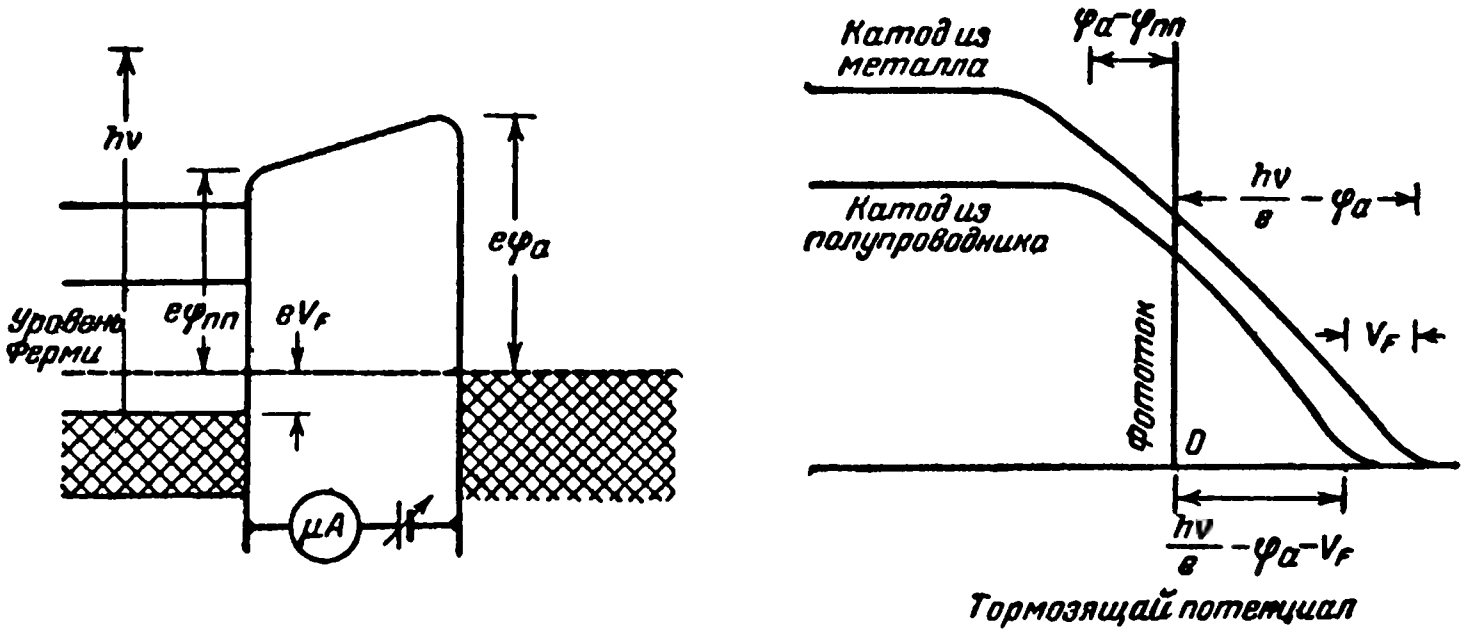


Рис. 22-18 Если металлический фотокатод на рис. 22-17 заменить полупроводником p-типа, от отсечка фототока наступит при тормозящем потенциале, отличающемся на величину V_F.

обусловленный тепловыми скоростями электронов). На рис. 22-17 справа показана типичная кривая зависимости фотоэлектронного тока от тормозящего потенциала. Если материал катода изменить, взяв металл с другой работой выхода, ток прекратится по-прежнему при напряжении $(h\nu/e) - \varphi_a$, если работа выхода катода меньше, чем у анода. Форма кривой различна у разных металлов, так как у них различны распределения по энергиям энергетических уровней в пределах частично заполненных зон.

Заменим теперь металл катода изучаемым полупроводником, причем все остальные условия опыта оставим неизменными. В полупроводнике заполнены все энергетические уровни в пределах валентной зоны, верхний край которой на eV_F ниже уровня Ферми на рис. 22-18. Поэтому прекращение фотоэлектронного тока будет иметь место не при $(h\nu/e) - \varphi_a$, а при $(h\nu/e) - \varphi_a - V_F$. Таким образом, на этом простом опыте можно независимо определить V_F, а также в принципе изменение его от температуры.

Если приложить тормозящее напряжение в противоположном направлении, чтобы частично компенсировать контактную разность потенциалов между этими двумя поверхностями, и постепенно увеличивать его, то фотоэлектронный ток будет расти, пока напряжение не достигнет величины $\varphi_а - \varphi_{пп}$. После этого ток останется постоянным, несмотря на дальнейшее увеличение напряжения, так как он ограничен работой выхода полупроводника $\varphi_{пп}$. При напряжении, при котором наступает такое

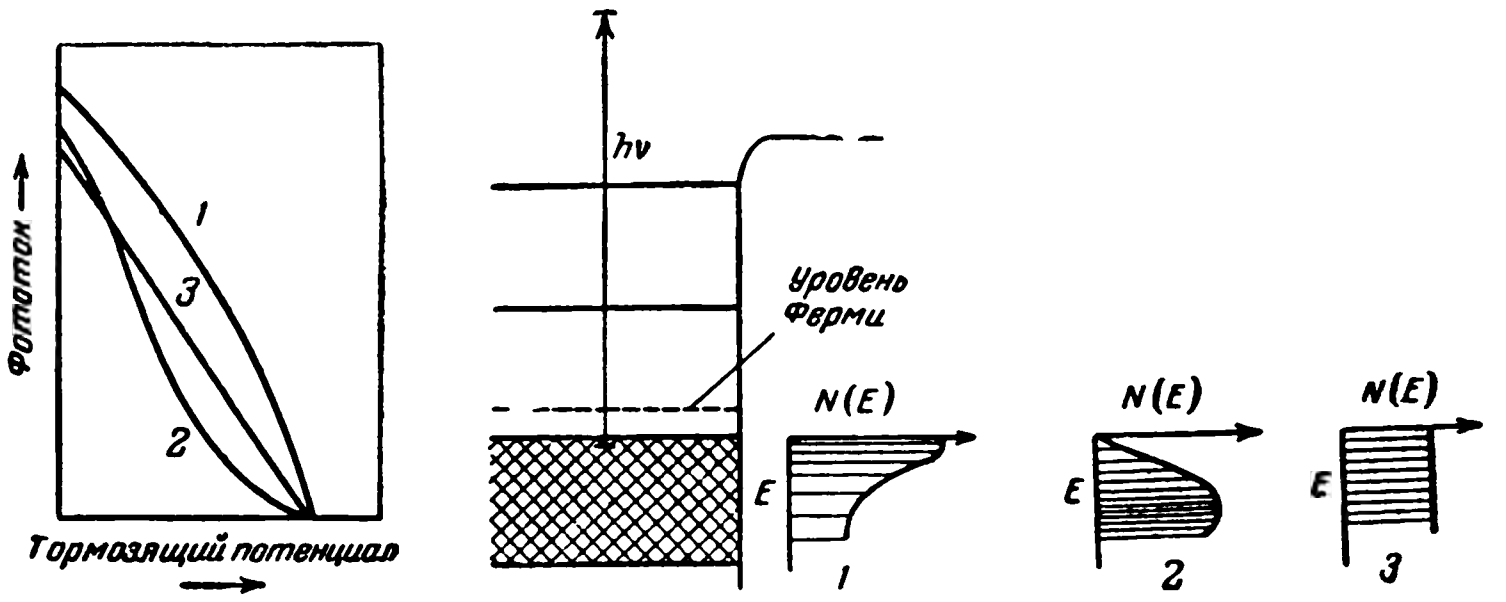


Рис. 22-19. Энергетическая плотность уровней электронов в валентной зоне полупроводника может быть определена из формы кривой тормозящего потенциала вблизи отсечки.

насыщение, $\varphi_{пп} = \varphi_а - U_{торм}$. Это явление показано на рис. 22-18. Таким образом, такое наблюдение является независимым способом измерения работы выхода поверхности полупроводника при условии, если известна работа выхода анода.

Можно также предположить, что из формы кривой фотоэлектронного тока в функции тормозящего потенциала вблизи прекращения тока можно определить функцию $N(E)$, определяющую плотность энергетических уровней в валентной зоне. На рис. 22-19 приведены три формы кривых фотоэлектронного тока, соответствующие трем типам распределения по энергиям энергетических уровней у вершины валентной зоны: 1) энергетические уровни плотнее у верхнего края зоны, чем в глубине ее, как можно предположить из рис. 15-11, который соответствует представлению о коллективизированных электронах; 2) расстояние между энергетическими уровнями больше и верхнего края зоны и меньше внутри нее; плотность энергетических уровней постоянна. Кривые фото-

электронного тока при тормозящем потенциале должны иметь форму, указанную на правой диаграмме, при условии, если все электроны в полупроводнике поглощают падающие фотоны с одинаковой вероятностью и если все электроны, имеющие после возбуждения достаточную энергию, чтобы выйти из поверхности катода, могут с одинаковой вероятностью выполнить это независимо от кинетической энергии, оставшейся у них после выхода.

К сожалению, картина не столь проста, так как обе эти вероятности зависят от энергии электрона, так что выводы о плотности распределения энергетических уровней в валентной зоне полупроводника гораздо сложнее. Однако такие выводы были сделаны для таких полупроводников [Л. 20], как теллур, греманий и бор, и для полуметаллов [Л. 21] — висмута, сурьмы и мышьяка. У полупроводников функция плотности распределения уровней по энергиям оказались типа *2* на рис. 22-19. Для полуметаллов распределения незакономерны и характеризуются довольно низкой плотностью в областях энергий, соответствующих вершинам распределений электронов в частично заполненных зонах. Следовательно, плохая проводимость этих материалов в сравнении с медью или серебром объясняется относительно малым количеством электронов в области энергий, в которой они способны участвовать в проведении тока.

Другим методом определения плотности распределения по энергиям электронных уровней в зонах твердых тел является изучение форм и структур рентгеновских линий, которые излучаются в результате электронных переходов между широкими валентными зонами твердых тел и узкими уровнями внутренних оболочек атомов. С помощью этого метода показано [Л. 22 и 23], что у разных твердых тел плотность уровней действительно выше в середине зоны, чем у ее краев. Эти открытия противоречат простой зонной теории, рассмотренной в гл. 15, что дает возможность лишний раз отметить, что если простая теория удовлетворительна для качественного описания явлений, то в ней имеются допущения, которые ограничивают ее количественное применение.

ЗАДАЧИ

1. Натриевая лампа равномерно излучает по всем направлениям монохроматический свет мощностью 20 *вт* с длиной волны 0,59 *мк*. Требуется доказать, что плотность фотонов на расстоянии 1 *м* от этой лампы равна $1{,}6 \cdot 10^4$ $1/см^3$.

2. По данным на рис. 22-4 докажите, что длина волны порога фотоэффекта для окиси меди равна примерно 0,77 *мк*.

3. Одна сторона пластинки из первоначально чистого германия толщиной 0,01 *см* равномерно облучается монохроматическим светом натрия с длиной волны 0,59 *мк* и интенсивностью 100 $мвт/см^2$. Пластинка достаточно тонка, так что рождающиеся на облучаемой поверхности электроны и дырки, диффундируя, распределяются в ней совершенно равномерно, прежде чем рекомбинируют. Время жизни электронов и дырок в этом образце и при данной концентрации частиц составляет 400 *мксек*. Предположим, что потеря на отражение света от поверхности равна 35% и что каждый фотон, проникающий в поверхность, производит одну пару частиц. До какой величины уменьшится удельное сопротивление окиси меди в результате облучения? (Ответ: около 1,5 *ом · см*.)

ЛИТЕРАТУРА

1. Becker J. A., Thermionic Electron Emission, Review of Modern Physics, 1935, vol. 7, p. 95.

2. Fan H. Y., Becker M., Infra-Red Absorption of Silicon, Physical Review, 1950, vol. 78, p. 178.

3. Jeans James, The Mathematical Theory of Electricity and Magnetism, Cambridge, 1908, chapter 18.

4. Briggs H. B., Infra-Red Absorption in Silicon, Physical Review, 1950, vol. 77, p. 727; Briggs H. B., Fletcher R. C., New Infra-Red Absorption Bands in p-Type Germanium, Physical Review, 1952, vol. 87, p. 1130.

5. Seitz F., Modern Theory of Solids, McGraw-Hill, New York, 1940, chapter 7.

Русский перевод: Мот Н. В., Черни Р., Электронные процессы в ионных кристаллах, Изд. иностранной литературы, 1950.

6. Mott N. F., Gurney R. W., Electronic Processes in Ionic Crystals, Oxford-Clarendon, 1940, chapter 3.

Русский перевод: Мот В., Черни Р., Электронные процессы в ионных кристаллах, Изд. иностранной литературы, 1950.

7. Smith Lloyd P., Report Sci. Ad., 21/49, 10 June, 1949.

8. Burstein E., Picus G., Sclar N., in Photoconductivity Conference, P. G. Breckenridge, Ed.: Wiley, New York, 1956.

9. Seitz F., Bulletin of the American Physical Society, 1951, vol. 26, p. 44.

10. Haynes J. R., Briggs H. B., Physical Review, 1952, vol. 86, p. 647.

11. Haynes J. R., New Radiation Resulting from Recombination of Holes and Electrons in Germanium, Physical Review, 1955, vol. 98, p. 1866.

12. Haynes J. R., Westphal W. C., Radiation Resulting from Recombination of Holes and Electrons in Silicon, Physical Review, 1956, vol. 101, p. 1676.

13. Schwarz E., New Photoconductive Cells, Nature, 1948, vol. 162, p. 614.

14. Gudden B., Pohl R., Über den zeitlichen Ansteig der Lichtelektrischen Leifähigkeit, Zeitschrift für Physik, 1921, vol. 6, p. 248.

15. James Hubert M., Conduction in Photoconductive PbS Films, Science, 1949, vol. 110, p. 254.

16. Rittner E. S., Concerning the Theory of Photoconductivity in Infra-Red-Sensitive Semiconducting Films, Nature, 1950, vol. 111, P. 685.

17. Sosnowski L. et al., Occurrence of Random Photovoltaic Barriers in Photoconductive Layers, Nature, 1947, vol. 160, p. 471.

18. Cornell Symposium on Solid Luminescent Materials, Wiley, F. R. Fonda and F. Seitz, Eds.

19—20. Apker L., Taft E., Dickley J., Photoelectric Emission and Contact Potentials of Semiconductors, Physical Review, 1948, vol. 74, p. 1462.

21. Apker L., Taft E., Dickley J., Some Semi-Metallic Characteristics of the Photoelectric Emission from As, Sb. and Bi, Physical Review, 1949, vol. 76, p. 270.

22. O'Bryan H. M., Skinner H. W. B., Characteristic X-Rays from Metals, Physical Review, 1934, vol. 45, p. 370.

23. Skinner H. W. B., The Soft X-Ray Spectroscopy of Solids, Philosophical Transactions of the Royal Society of London, 1940, vol. 239A, p. 95.

24. Гуро Г. М., К теории фотогальваномагнитного эффекта в полупроводниках, ЖТФ, 1958, 28, 5, 1036—1045.

ГЛАВА ДВАДЦАТЬ ТРЕТЬЯ

ТЕРМОЭЛЕКТРИЧЕСКИЕ ЭФФЕКТЫ В ПОЛУПРОВОДНИКАХ

Физика термоэлектрического эффекта рассмотрена в гл. 2, где кратко описаны его особенности. В настоящей главе мы исследуем происхождение термоэлектрического напряжения с точки зрения физики полупроводников, а также обсудим, какие основные свойства полупроводников можно установить посредством термоэлектрических измерений.

В 1826 г. Зеебек впервые сообщил, что в цепи, содержащей два различных металла с разными температурами контактов, имеет место термоэлектрический эффект.

Поэтому термоэлектрический эффект часто называется эффектом Зеебека. Он лежит в основе метода измерения температуры, при котором один из контактов имеет эталонную температуру (обычно температура тающего льда), а другой вводится в среду, температуру которой надо определить. Термометрия, применяющая термо-

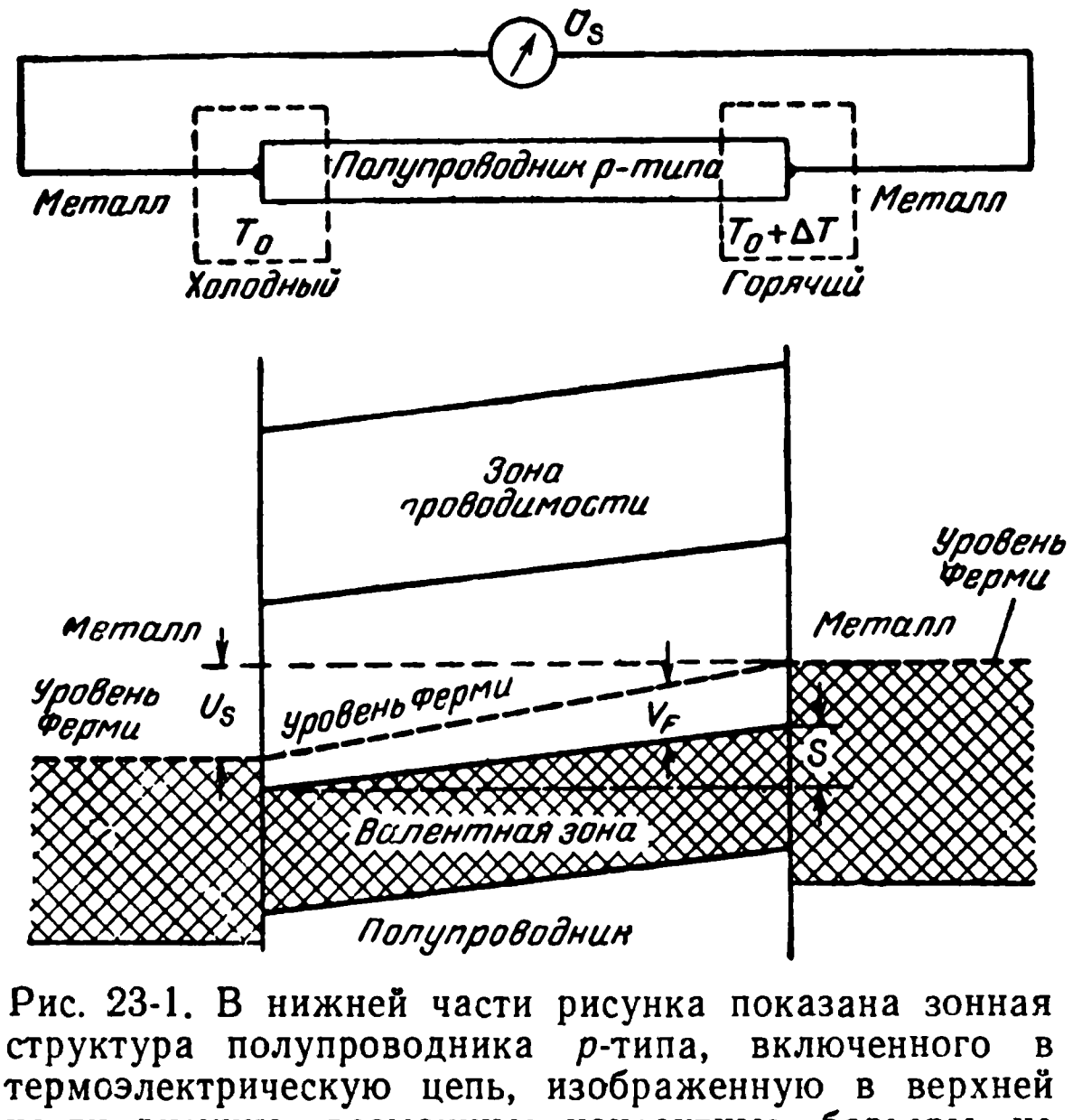


Рис. 23-1. В нижней части рисунка показана зонная структура полупроводника *p*-типа, включенного в термоэлектрическую цепь, изображенную в верхней части рисунка; возможные контактные барьеры не учитываются. Энергии на этом рисунке измеряются эквивалентными напряжениями.

пары из высокостабильных металлов или сплавов, стала очень важной частью технологии промышленных измерений и контроля.

Полупроводники в цепях с металлами также обнаруживают термоэлектрический эффект. Кроме того, величины термоэлектрических коэффициентов полупроводников часто гораздо больше, чем соответствующие величины для металлов. Далее, термоэлектрические свойства полупроводников могут быть использованы для получения сведений о положении уровня Ферми по отношению к краям зон и о концентрации подвижных частиц.

Типичная термоэлектрическая цепь с полупроводни-

ком показана на рис. 23-1. Полупроводник имеет форму длинного тонкого прутка с металлическими контактами на обоих концах. Оба контакта сделаны из одинакового металла, как и обе проволоки, ведущие к вольтметру, измеряющему термо-э. д. с. Оба зажима этого прибора должны иметь одинаковую, например комнатную, температуру, в противном случае в цепи создается нежелательная дополнительная термо-э. д. с. Один конец прутка (обычно холодный) все время имеет эталонную температуру T_0, в то время как другой конец помещается в среду при температуре $T_0+\Delta T$. Будем считать градиент температуры вдоль прутка постоянным, а материал прутка p-типа. При этих условиях уровень Ферми в металле и полупроводнике и края зон полупроводника устанавливаются так, как показано в нижней части рис. 23-1. Края зон полупроводника становятся наклонными, как изображено на рисунке, а уровень Ферми тоже наклоняется, но несколько круче. Уровень Ферми в полупроводнике является продолжением уровней Ферми в металлах обоих контактов [1]. Уровни Ферми в металлических проволоках, идущих от контактов, не меняются при изменении температуры, как это происходит из-за температурного градиента с уровнем Ферми внутри полупроводника; поэтому на рис. 23-1 уровни Ферми в металлических проводниках показаны горизонтальными и термо-э. д. с. в цепи в основном определяется разностью высот уровня Ферми между концами полупроводника. В случае однородного полупроводника p-типа, приведенного в качестве примера на рис. 23-1, правый конец прутка приобретает отрицательный потенциал относительно левого конца. Физические причины перекоса энергетических зон и уровней Ферми мы рассмотрим позднее. Для наглядности наклон энергетических уровней на рисунке сильно преувеличен.

Если разность температур мала, то характеристики, изображенные на рис. 23-1, зависят линейно от температуры, а следовательно, также линейно зависят от расстояния по горизонтали на диаграмме. Коэффициент пропорциональности между U_s и ΔT называется «удельной термо-э. д. с.» (термо-э. д. с. на 1° С); его физическая размерность вольты на градус, и он далее обозна-

---

[1] Возможность наличия выпрямляющих барьеров в этих контактах никак не влияет на рассуждения.

чен символом Q, так что $Q=U_s/\Delta T$. Если одновременно повышать или понижать температуру горячего и холодного спаев, не изменяя при этом ΔT, то окажется, что Q зависит от средней температуры. Изменения Q от T можно также наблюдать, сохраняя холодный спай при температуре T_0 и графически изображая отношение термо-э. д. с. U_s в зависимости от изменения температуры $T_0+\Delta T$ горячего спая. Наклон этой кривой при любой температуре T определяет величину Q при этой температуре.

Наклон энергетических зон и уровня Ферми в полупроводниках при наличии температурного градиента можно объяснить следующим образом. Предположим, что температуры T и $T_0+\Delta T$ лежат в области температур, в которой свойства полупроводника определяются примесями в нем, и допустим, что он активирован равномерно. Тогда концентрация основных носителей (положительных дырок) у правого конца образца на рис. 23-1 выше, чем у левого. Следовательно, будет иметь место диффузия положительных дырок под влиянием температурного градиента справа налево. Поэтому холодный конец образца становится положительно заряженным и между двумя концами прутка появляется такая разность потенциалов, что обратный ток, обусловленный этой разностью потенциалов, уравновешивает диффузионный поток. Эта разность потенциалов определяется расстоянием S по вертикали на диаграмме.

Даже если не существует никакого градиента концентрации основных носителей (как будет в том случае, если температуры T_0 и $T_0+\Delta T$ лежат в таком диапазоне, что акцепторы все насыщены), все же между концами образца появляется небольшая разность потенциалов. Эта разность потенциалов обусловлена тем фактом, что дырки, пересекающие воображаемое поперечное сечение образца вследствие беспорядочного теплового движения, движутся несколько быстрее справа налево, чем в обратном направлении, так как они двужутся от более теплого конца образца и имеют более высокие тепловые кинетические энергии. Однако в проводниках со свойствами, зависящими от примесей, и в собственных полупроводниках этот эффект незначителен по сравнению с влиянием диффузии и им в приближенной теории можно пренебречь.

23-1. ВЫЧИСЛЕНИЕ ТЕРМО-Э. Д. С. И УДЕЛЬНОЙ ТЕРМО-Э. Д. С.

Интересно напомнить, что, как показано в § 19-4, разность потенциалов появляется и между концами полупроводника с градиентом концентрации активатора, даже если все тело находится в термическом равновесии при постоянной температуре. В рассмотренном там случае уровень Ферми соответствовал постоянному значению энергии во всем теле. В настоящем случае, однако, при наличии разности температур между концами образца стержень не находится в термодинамическом равновесии и нет никаких оснований для того, чтобы энергия Ферми сохранялась постоянной. Фактически она и не остается постоянной, что видно из рис. 23-1. В гл. 17 мы пришли к выводу, что уровень Ферми в полупроводниках p-типа с повышением температуры приближается к середине энергетической щели. Поэтому общее повышение уровня Ферми на рис. 23-1 от холодного конца образца к теплому концу определяется суммой подъема верхнего уровня валентной зоны плюс увеличение, обусловленное тем, что уровень Ферми поднимается относительно валентной зоны. Другими словами,

$$U_s = S + \frac{dV_F}{dT}\Delta T. \qquad (23\text{-}1)$$

Таково напряжение, показываемое высокоомным вольтметром в металлической части цепи.

Вычисление термоэлектрической величины U_s в терминах основных физических величин просто для случаев простых полупроводников, в которых надо принимать во внимание только один тип проводящих частиц. В образце p-типа, представленном на рис. 23-1, плотность тока, идущего справа налево вследствие диффузии, выражается через

$$I_{R\to L} = D_p e \frac{dp}{dx}.$$

Обратный ток слева направо, обусловленный электрическим полем, возникшим в результате этого явления, выражается как

$$I_{L\to R} = p\mu_p e E_x.$$

В установившемся режиме эти два тока равны, что дает:

$$D_p e \frac{dp}{dx} = p\mu_p e E_x. \qquad (23\text{-}2)$$

Вместо D_p можно написать $kT\mu_p/e$ из соотношения Эйнштейна; предположив, что электрическое поле однородно (как это имеет место при небольшой разности температур и постоянном градиенте), мы можем заменить E_x через S/L, где L — длина образца. С этими подстановками после некоторых преобразований уравнение (23-2) дает:

$$\frac{dp}{dx} = \frac{pe}{kT} \cdot \frac{S}{L}. \qquad (23\text{-}3)$$

Однако

$$\frac{dp}{dx} = \frac{dp}{dT} \cdot \frac{dT}{dx} = \frac{dp}{dT} \cdot \frac{\Delta T}{L},$$

так что уравнение (23-3) превращается в

$$\frac{dp}{dT} = \frac{pe}{kT} \cdot \frac{S}{\Delta T}. \qquad (23\text{-}4)$$

Согласно статистике Ферми

$$p = 2UT^{3/2} e^{-eV_F/kT}.$$

Дифференцируя по T и подставляя результат в уравнение (23-4), после упрощений получаем:

$$S = \left(-\frac{dV_F}{dT} + \frac{V_F}{T} + \frac{3}{2} \cdot \frac{k}{e} \right) \Delta T. \qquad (23\text{-}5)$$

Комбинируя уравнения (23-5) и (23-1), имеем:

$$U_s = \left(\frac{V_F}{T} + \frac{3}{2} \cdot \frac{k}{e} \right) \Delta T, \qquad (23\text{-}6)$$

или, так как $Q = U_s/\Delta T$,

$$Q = \frac{V_F}{T} + \frac{3}{2} \cdot \frac{k}{e}. \qquad (23\text{-}7)$$

Два последних уравнения для термо-э. д. с. и для удельной термо-э. д. с. являются только приблизитель-

ными. При их выводе не учтен описанный ранее эффект, в соответствии с которым небольшой вклад в S происходит также от разности средних тепловых энергий частиц, диффундирующих в образце в противоположных направлениях. С поправкой на этот эффект получаем для U_s и Q:

$$U_s = \left(\frac{V_F}{T} + 2\,\frac{k}{e}\right)\,\Delta T; \tag{23-8}$$

$$Q = \frac{V_F}{T} + 2\,\frac{k}{e}\,. \tag{23-9}$$

Подобный вывод справедлив и для полупроводников n-типа; он приводит к уравнениям, сходным с двумя приведенными выше, за исключением того, что знаки у U_s и Q отрицательны и вместо V_F входит выражение $V_g - V_F$.

Уравнение (23-9) дает нам независимый метод определения V_F в полупроводнике p-типа. Зная его, нетрудно вычислить концентрацию дырок p.

Поскольку

$$p \approx 2UT^{3/2}e^{-eV_F/kT},$$

то, следовательно, из уравнения (23-9)

$$p = 2UT^{3/2}e^{-eQ/k+2}. \tag{23-10}$$

Необоходимо помнить, что эти уравнения применимы только для температур ниже той, при которой полупроводник становится компенсированным. Для более высоких температур вывод должен учитывать одновременно оба сорта носителей, и получающиеся в результате уравнения становятся более сложными, например:

$$Q = \frac{\frac{1}{T}\left[p\mu_p\left(V_F + 2\,\frac{kT}{e}\right) - n\mu_n\left(V_g - V_F + 2\,\frac{kT}{e}\right)\right]}{p\mu_p + n\mu_n}\,. \tag{23-11}$$

23-2. ЭФФЕКТЫ ПЕЛЬТЬЕ И ТОМСОНА

В 1834 г. Пельтье наблюдал, что если через контакт двух различных металлов проходит ток, контакт нагревается или охлаждается в зависимости от направления тока. Этот нагрев больше обычного, вызываемого джоу-

левым теплом, которое обусловлено рассеянием мощности в сопротивлениях контакта и присоединенных к нему проводников. Степень нагрева или охлаждения выражается коэффициентом π, который равен числу джоулей тепла, освобожденных или поглощенных в контакте, на кулон прошедшего через него заряда. Следовательно, коэффициент π имеет физическую размерность напряжения, возникающего на контакте. Величина коэффициента π различна в зависимости от рода материалов,

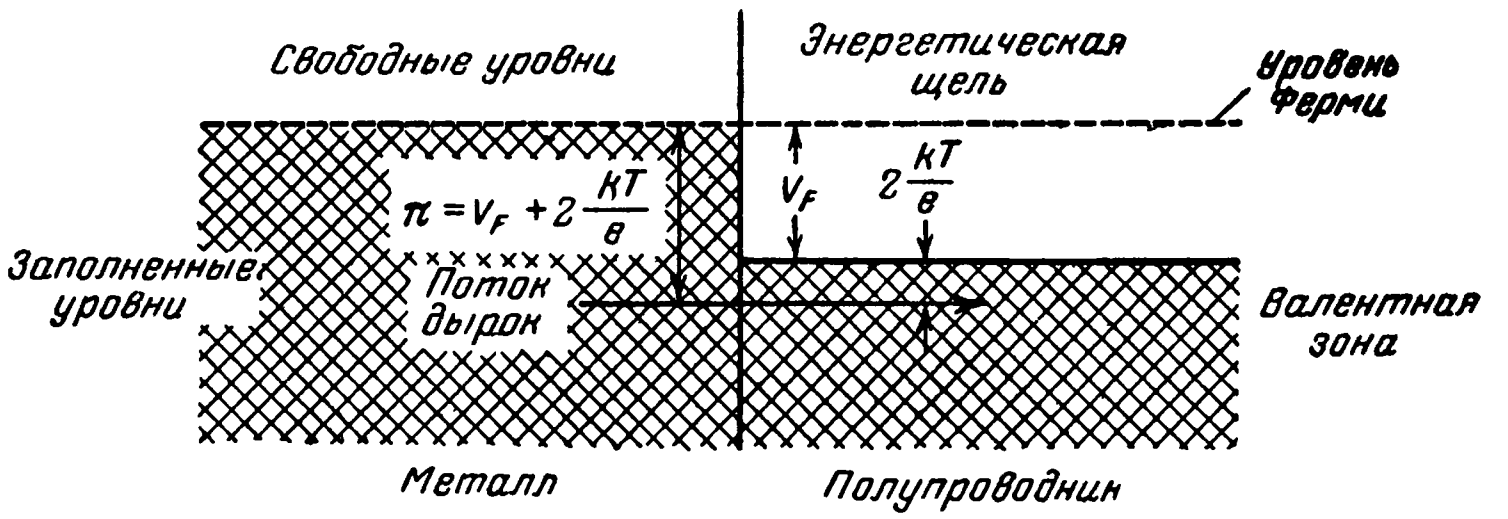


Рис. 23-2 В контакте такого рода происходит охлаждение Пельтье, если поток дырок идет слева направо

находящихся в контакте, причем обычно для контакта металла с полупроводником эта величина больше, чем для контакта двух металлов.

Мы знаем, что нагрев и охлаждение Пельтье обусловлены тем фактом, что когда носители заряда проходят через контакт, их средняя кинетическая энергия меняется. Механизм этого изменения можно объяснить, исходя из энергетической диаграммы на рис. 23-2. На ней показан контакт металла и полупроводника *p*-типа. В этом примере не учитывается возможное наличие выпрямляющего барьера, так как учет последнего не изменяет приведенного доказательства. Уровень Ферми в металле, совпадающий с вершиной распределения электронов в частично заполненной зоне, продолжается уровнем Ферми в полупроводнике. Пусть теперь положительные дырки идут через контакт слева направо. В металле средняя тепловая энергия дырки отсчитывается вниз от уровня Ферми. Для того чтобы пройти через контакт в полупроводник, дырка должна иметь энергию, по меньшей мере на V_F бóльшую, чем ее средняя энергия в металле. Следовательно, только часть общего количества

дырок в металле имеет достаточную энергию, чтобы перейти через контакт. Каждая дырка, пересекающая контакт, приносит в полупроводник в среднем тепловую энергию $2kT$ *. Следовательно, каждая дырка, проходящая контакт, забирает из металла энергию eV_F+2kT. Напряжения, эквивалентные этим величинам, приведены на диаграмме рис. 23-2. Любопытно, что эта энергия не сообщается полупроводнику, так как эти дырки, после того как они попали в полупроводник, находятся в термическом равновесии с остальными имеющимися там дырками. Энергия просто непрерывно поглощается в приконтактной области, вследствие чего контакт охлаждается. Металл должен непрерывно снабжать дырки энергией, чтобы восстановить распределение энергии дырок, нарушенное переходом в полупроводник дырок с более высокой энергией. Эта восстанавливающая энергия получается за счет тепловой энергии решетки металла, в результате чего и происходит охлаждение.

Ясно, что если это рассуждение провести в обратном порядке, т. е. для такого же тока, идущего в обратном направлении, то в металле должно выделиться такое же количество тепла на расстоянии в несколько средних свободных пробегов от перехода. Теперь можно ответить на вопрос, куда исчезает тепло, отбираемое у металла в том случае, когда дырки движутся слева направо. Это то тепло, которое передается другим участкам цепи, причем общее количество тепла в цепи не изменяется, если пренебречь выделением джоулева тепла.

Таким образом, напряжение нагрева или охлаждения в контакте между металлом и полупроводником p-типа выражается следующим образом:

$$\pi = \pm \left(V_F + \frac{2kT}{e} \right) \text{ дж/к}, \qquad (23\text{-}12)$$

причем знак зависит от направления тока.

В цепи термопары на рис. 23-1 π различно для двух контактов с различными температурами, так как в урав-

---

* Можно показать, что в явлениях переноса подобного рода средняя кинетическая энергия, переносимая одной частицей, равна $2\,kT$, а не $^3/_2\,kT$, как получается при определении средней кинетической энергии частицы в равновесии.

нении (23-12) T и V_F (которое само является функцией T) различны для обоих контактов.

В начале этой главы описано возникновение электрического поля в равномерно активированном полупроводнике при наличии градиента температуры. Разность потенциалов между концами образца, между которыми поддерживается разность температур ΔT, была обозначена S (рис. 23-1). Если ток положительных дырок идет слева направо по полупроводнику на рис. 23-1, то положительные дырки, проходя вдоль образца, теряют потенциальную энергию eS на каждую дырку. В то же время они приобретают кинетическую энергию переноса в среднем $2k\Delta T$ на дырку, так как они переходят от холодного конца стержня к теплому концу. Сумма этих двух величин — $eS+2\,k\Delta T$ составляет чистый прирост энергии на каждую дырку при прохождении всей длины стержня. Это увеличение энергии может быать положительным или отрицательным в зависимости от того, какое слагаемое больше. В результате энергия либо извлекается из кристаллической решетки по всей длине полупроводника, либо выделяется в ней, что вызывает либо охлаждение, либо нагрев образца.

Это явление предсказал более 100 лет назад В. Томсон (впоследствии лорд Кельвин). В 1855 г. он из чисто термодинамических соображений пришел к заключению, что проводник, по которому проходит электрический ток и в котором существует продольный градиент температуры, должен нагреваться или охлаждаться в зависимости от направления тока, помимо и независимо от необратимой потери в нем джоулевого тепла. Этот вывод был затем проверен для ряда веществ задолго до того, как были объяснены физика твердого тела и механизм этого явления. Тепло, поглощаемое и отдаваемое проводящим образцом в джоулях на кулон и на единицу температурного градиента, называется поэтому коэффициентом Томсона Θ. Он считается положительным, если положительный ток (т. е. идущий в направлении положительного температурного градиента) увеличивает среднюю величину полной энергии носителей (охлаждает образец). Таким образом, коэффициент Θ имеет физическую размерность вольты на градус и зависит от вещества, из которого сделан образец. Исходя из определения коэффициента Θ, а также из выводов последнего пара-

графа, касающегося общей энергии, полученной или потерянной на 1 $к$ положительных дырок, движущихся в направлении разности температур ΔT в полупрсводнике p-типа, мы можем написать:

$$\theta\Delta T = -S + \frac{2k\Delta T}{e} \; дж/к. \qquad (23\text{-}13)$$

23-3. СООТНОШЕНИЕ МЕЖДУ ТЕРМО-Э. Д. С. И КОЭФФИЦИЕНТАМИ π И θ

В этом параграфе мы рассмотрим эффекты Зеебека, Пельтье и Томсона в замкнутой термоэлектрической це-

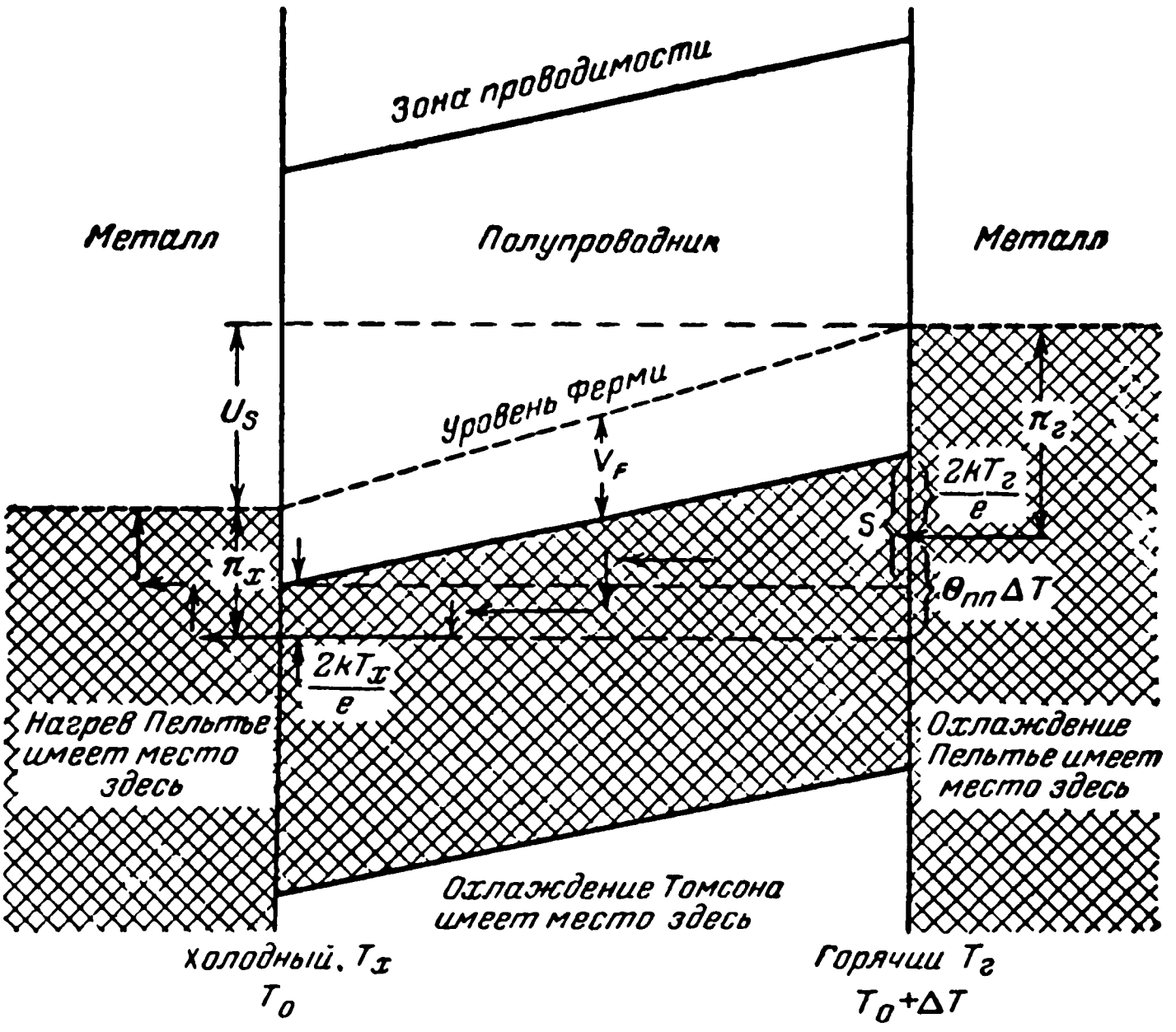


Рис 23-3 На этой энергетической диаграмме отложены эквивалентные напряжения термоэлектрической цепи с полупроводником p-типа. Горизонтальные стрелки показывают нормальный поток дырок.

пи (рис. 23-3). Мы покажем, что эти три эффекта связаны и приведем некоторые следствия из существующего между ними соотношения.

Так как коэффициенты π и Θ в уравнениях (23-12) и (23-13) выражены в джоулях на кулон, они опреде-

ляют скачки напряжения, которые имеют место в контактах, для коэффициента π или разность потенциалов в направлении температурного градиента в полупроводнике и металлической части цепи для коэффициента Θ. Как читатель, вероятно, уже догадался, термо-э. д. с., возникающая в замкнутой термоэлектрической цепи, является суммой четырех отдельных напряжений:

1) напряжение Пельтье на горячих контактах минус;

2) напряжение Пельтье на холодных контактах плюс;

3) разность напряжения Томсона между концами металлической части цепи минус;

4) разность напряжения Томсона между концами полупроводника.

Таким образом,

$$U_s = \pi_г - \pi_х + \theta_м \Delta T - \theta_{пп} \Delta T, \qquad (23\text{-}14)$$

где индексы г, х, м, пп соответственно обозначают: горячий, холодный, металл и полупроводник. Так как θ непостоянно для всех значений T, члены $\theta \Delta T$ должны быть заменены интегралами вида

$$\int_{T_х}^{T_г} \theta dT,$$

если ΔT превышает несколько десятков градусов. Для большинства металлов $\theta_м$ очень мало по сравнению с обычными значениями $\theta_{пп}$. Поэтому третий член уравнения (23-14) можно в дальнейшем отбросить. Этим оправдывается то, что уровни Ферми представлены горизонталями для металлических частей цепи, присоединенных к горячим и холодным контактам, а также утверждение, что термо-э. д. с. равна просто вертикальному смещению уровней Ферми металлических контактов на концах диаграммы, приведенной на рис. 23-3. Прямой подстановкой в приведенном выше уравнении соответствующего выражения для $\theta_{пп} \Delta T$ из уравнения (23-13) и выражения для S из уравнения (23-5) получаем для U_s тот же результат, который дан в уравнении (23-6), чем доказывается, что представления, на основании которых мы установили связь между термо-э. д. с. и коэффициентами Пельтье и Томсона, совершенно последовательны.

Если мы замкнем термоэлектрическую цепь через вольтметр с большим сопротивлением (а не через вольтметр с бесконечным сопротивлением), то пока будет поддерживаться разность температур ΔT, будет течь небольшой ток. Его величина получится в результате деления U_s на сумму сопротивлений полупроводника и внешней цепи. Направление этого тока на диаграмме, показанной на рис. 23-3, определяется движением дырок в полупроводнике от горячего к холодному концу и электронов, перемещающихся от горячего к холодному концу в остальной части цепи. В результате такого перемещения тепло накапливается в одних частях цепи и теряется в других. Если ток равен 1 *а*, мы можем сделать следующие количественные заключения. Когда дырки входят в полупроводник из горячего контакта, тепло поглощается со скоростью $I_{\pi г}$ *вт*. Это тепло должно выделяться с той же быстротой из источника тепла, поддерживающего постоянную температуру $T_0+\Delta T$ горячего контакта. В полупроводнике тепло выделяется в решетку полупроводника или поглощается из нее в зависимости от того, является ли $\Theta_{пп}$ положительным или отрицательным. Эта скорость выделения или поглощения равна $I\Theta_{пп}\Delta T$ *вт*, и это тепло должно рассеиваться в окружающей среде или получаться из нее или же равновесие сохраняется за счет теплопроводности через полупроводник между резервуарами холода и тепла. Наконец, тепло приносится в металл током на холодном контакте с быстротой $I_{\pi x}$ и это тепло надо отвести в резервуар холода, чтобы сохранить на контакте температуру T_0. Разность между теплом, входящим в систему и выходящим из нее в секунду, составляет $I(\pi_г+ +\Theta_{пп}\Delta T-\pi_x)$ *вт*. Именно это количество тепла, поглощенного системой, дает энергию для поддержания термотока, и это тепло в конечном счете рассеивается в виде джоулевого тепла во всех частях цепи.

Если в цепь включена батарея достаточно высокого напряжения и такой полярности, которая заставит ток течь в обратном направлении, все тепловые явления обратятся, кроме джоулева тепла, которое зависит от I^2 и поэтому необратимо. Если вначале вся цепь имела постоянную температуру, то ток нагреет некоторые ее части, а другие охладит; в частности, левый контакт охладится, а правый нагреется.

Следует напомнить, что проведенные выше рассуждения относились к полупроводнику p-типа, температура которого ниже температуры достижения компенсированного состояния. Если эти соображения применить к электронам вместо дырок, то можно получить сведения о поведении полупроводника n-типа. При более высоких температурах необходимо учитывать одновременное присутствие электронов и дырок. В результате получаются более сложные закономерности, обсуждение которых не входит в нашу задачу.

23-4. ДРУГИЕ ПОЛЕЗНЫЕ ТЕРМОЭЛЕКТРИЧЕСКИЕ СООТНОШЕНИЯ

Выведем несколько дополнительных соотношений между различными термоэлектрическими величинами. Хотя они и не объясняют физические явления и не углубляют теорию, тем не менее они позволяют извлечь максимальные сведения из тех данных, которые на первый взгляд кажутся совершенно недостаточными.

Сравнение Q из уравнения (23-9) с π из уравнения (23-12) сразу показывает, что

$$\pi = QT. \tag{23-15}$$

Преобразовав уравнения (23-15), (23-5)*, (23-12) и (23-13), можно получить:

$$\theta_{\text{пп}} = T\frac{dQ}{dT}. \tag{23-16}$$

Уравнения (23-15) и (23-16) были выведены Томсоном 100 лет назад чисто термодинамическим путем [Л. 1]. Поэтому они называются соотношениями Томсона (или Кельвина). Используя эти соотношения и определение $Q = dU_э/dT$, мы можем написать следующие равенства:

$$\frac{dU_э}{dT} \equiv Q = \frac{\pi}{T} = \int_0^T \frac{\theta_{\text{пп}}}{T}\, dT. \tag{23-17}$$

---

* Уравнение (23-5) должно быть исправлено так, чтобы коэффициент у члена k/e равнялся 2, а не $^3/_2$ [см § 23-2, идущий за уравнением (23-7)].

В этом выражении U_s, Q, π и $\Theta_{пп}$ являются функциями температуры. В частности, если мы можем экспериментально определить график зависимости какой-либо из этих величин от T, то мы сможем получить графики и всех других величин как функции T с помощью уравнения (23-17).

Чтобы показать, насколько полезны термоэлектрические измерения для объяснения различных сторон поведения полупроводников, рассчитаем графики зависимо-

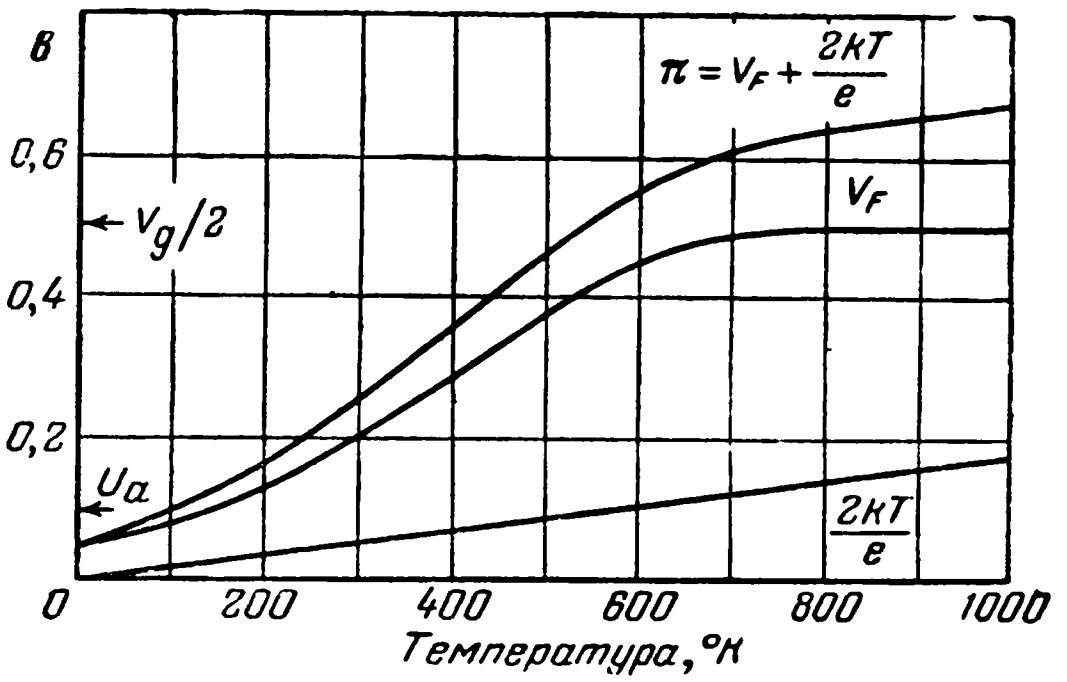


Рис. 23-4. Для данного образца полупроводника p-типа кривая коэффициента Пельтье является суммой кривых U_f и $2kT/e$ по уравнению (23-12). $V_g = 1{,}0$ *в*, $N_a = 10^{16}$ *см*^{3}; $U_a = 0{,}1$ *в*.

сти величин U_s, Q, π и $\Theta_{пп}$ от температуры для гипотетического полупроводника и затем сравним их с кривыми, полученными экспериментально для реального полупроводника.

За основу наших вычислений возьмем воображаемый полупроводник, который имеет ширину энергетической щели точно 1 *эв*. Мы будем считать этот образец полупроводником p-типа, имеющим равномерно распределенную концентрацию акцепторов N_a в 10^{16} *см*$^{-3}$. Уровни энергии акцепторов лежат на 0,1 *эв* выше вершины валентной зоны в энергетической схеме полупроводника.

Расчет лучше всего произвести, определяя вначале положение уровня Ферми как функцию температуры. Если мы знаем зависимость V_F от T, мы можем построить график зависимости π от T по уравнению (23-12), а затем, используя соотношения (23-17), вычислим или графически изобразим все остальное.

На рис. 23-4 приведен график зависимости V_F от T

для данного образца полупроводника. Диаграмма вычислена с помощью статистики Ферми, приведенной в гл. 17. На рис. 23-4 приведен также график величины $2\,kT/e$. И, наконец, там же приведен график зависимости π от T, который согласно уравнению (23-12) является просто суммой двух других графиков. Для данного образца уровень Ферми пересекает энергетический уро-

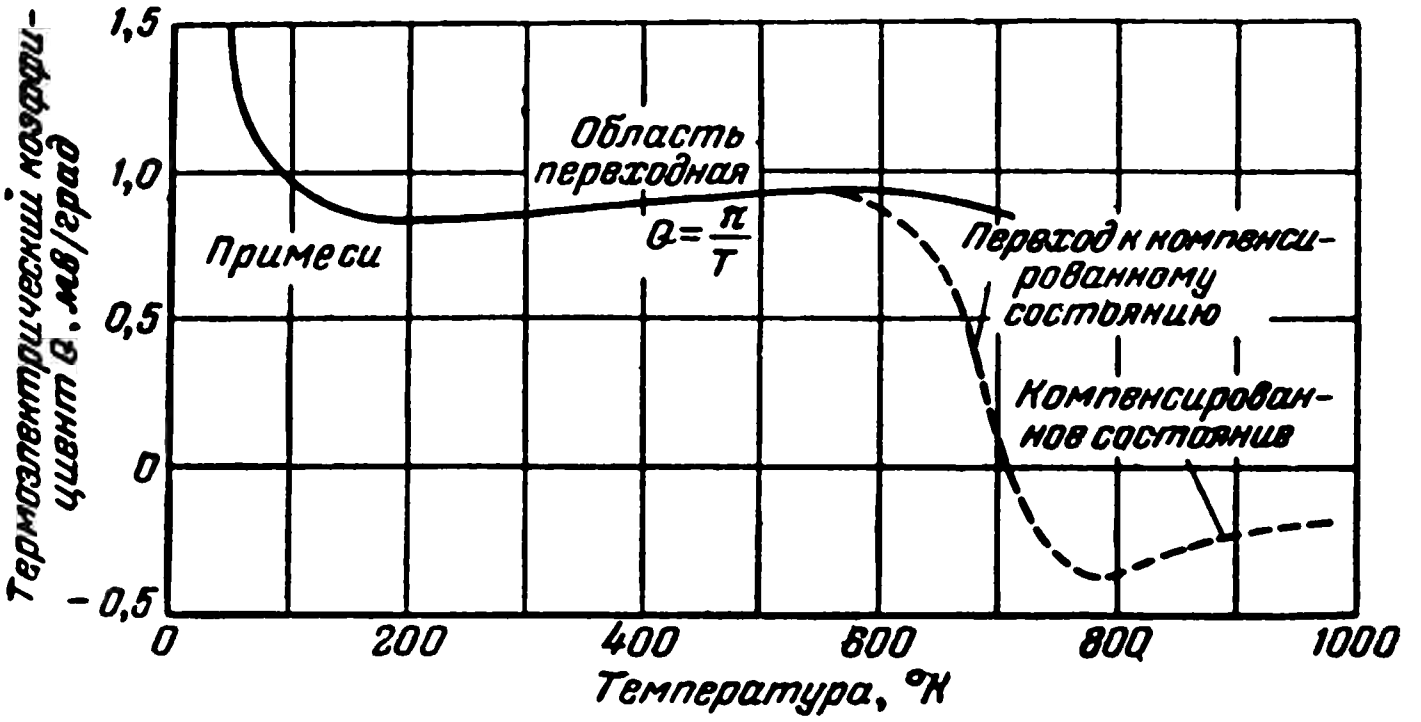


Рис. 23-5. Кривая термоэлектрического коэффициента, построенная по графику коэффициента Пельтье из рис 23-4 с помощью уравнения (23-17).

вень акцепторов примерно при 150° С. Начиная с несколько большей температуры (скажем, с 200° С), акцепторы приближаются к состоянию насыщения, и начинается большая переходная область температур. Верхняя граница этой переходной области простирается до 600° С, когда уровень Ферми становится примерно на kT ниже середины энергетической щели.

На рис. 23-5 сплошной кривой изображена зависимость Q от температуры, полученная из кривой для π на предыдущем рисунке. Пунктирная кривая изображает ожидаемое поведение этого полупроводника при приближении и переходе в температурную область компенсированного состояния, где простая теория становится непригодной, и надо учитывать присутствие электронов в концентрации, сравнимой с концентрацией дырок. В частности, если подвижность электронов больше подвижности дырок, то кривая Q пересечет ось и перейдет в отрицательную область. При еще более высоких температурах подвижность как электронов, так и дырок

уменьшается из-за влияния тепловых колебаний решетки и Q уменьшается по абсолютной величине, хотя все еще остается отрицательным.

На рис. 23-6 приведена теоретическая кривая для термо-э. д. с. как функции температуры для данного образца полупроводника Эта кривая получена интегри-

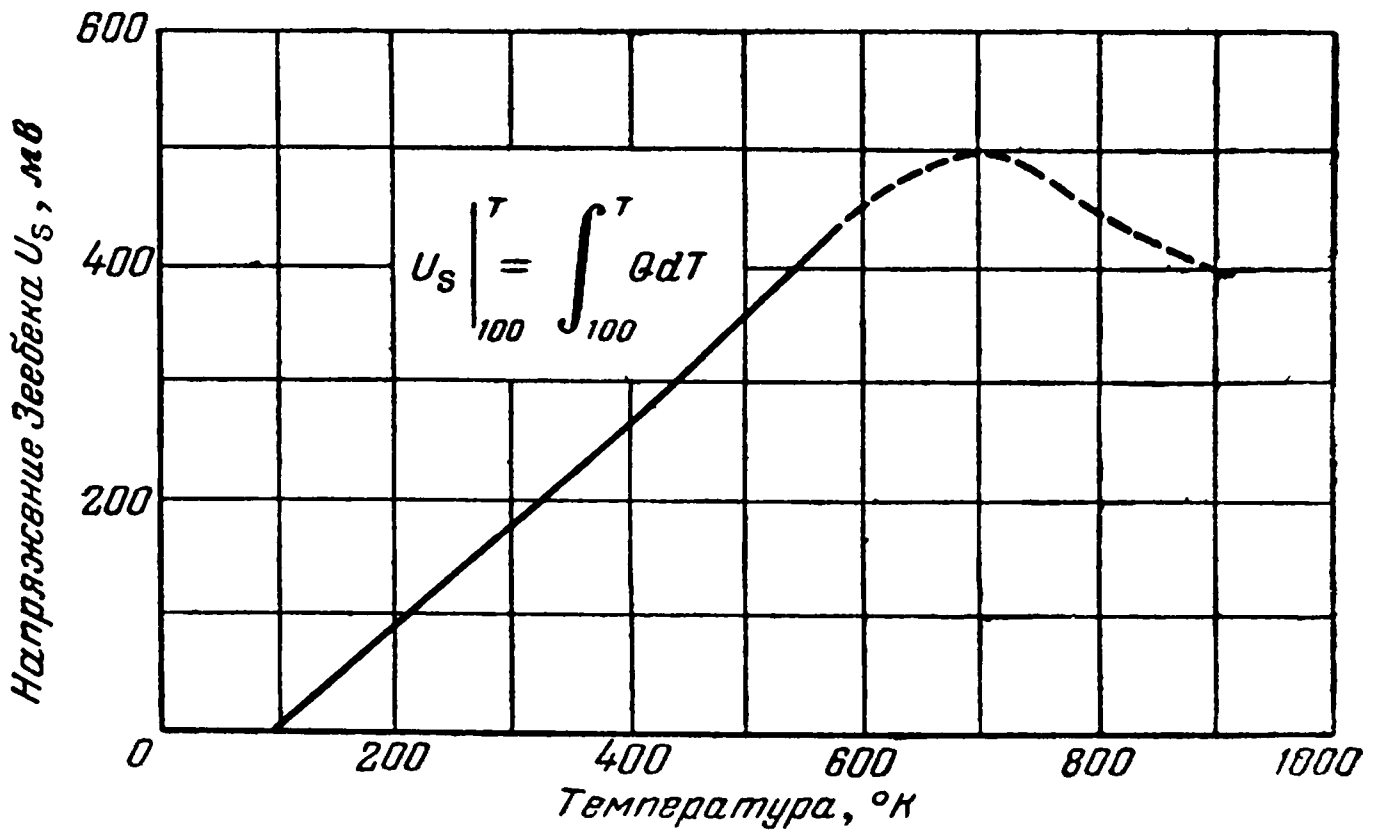


Рис 23-6. Кривая напряжения Зеебека в зависимости от температуры для термопары из полупроводника p-типа с холодным контактом при 100° К, построенная по кривой Q на рис. 15-5 на основании уравнения (23-17).

рованием величины Q из рис. 23-5 в соответствии с уравнением (23-17). При расчетах предположено, что холодный контакт имеет температуру 100° С, и поэтому эта температура принималась в качестве нижнего предела при интегрировании. Поведение в области компенсированного состояния изображено пунктирной кривой. Наконец, рис. 23-7 дает теоретическую кривую для коэффициента Томсона как функции температуры, полученную после выполнения действий, указанных в уравнении (23-17).

Следует также заметить, что если можно получить экспериментально любую из кривых, изображенных на рис. 23-4—23-7, то остальные могут быть построены по ней. Обычно измеряют термо-э. д. с. в функции температуры. Знак термо-э. д. с. в области низких температур обычно сразу указывает на тип проводимости полупроводника. Из кривой зависимости π от температуры мож-

но видеть, каково поведение уровня Ферми в функции температуры, и отсюда получить ширину энергетической щели, концентрацию основных носителей при любой температуре в области примесного состояния или в переходной области, концентрацию эффективных доноров

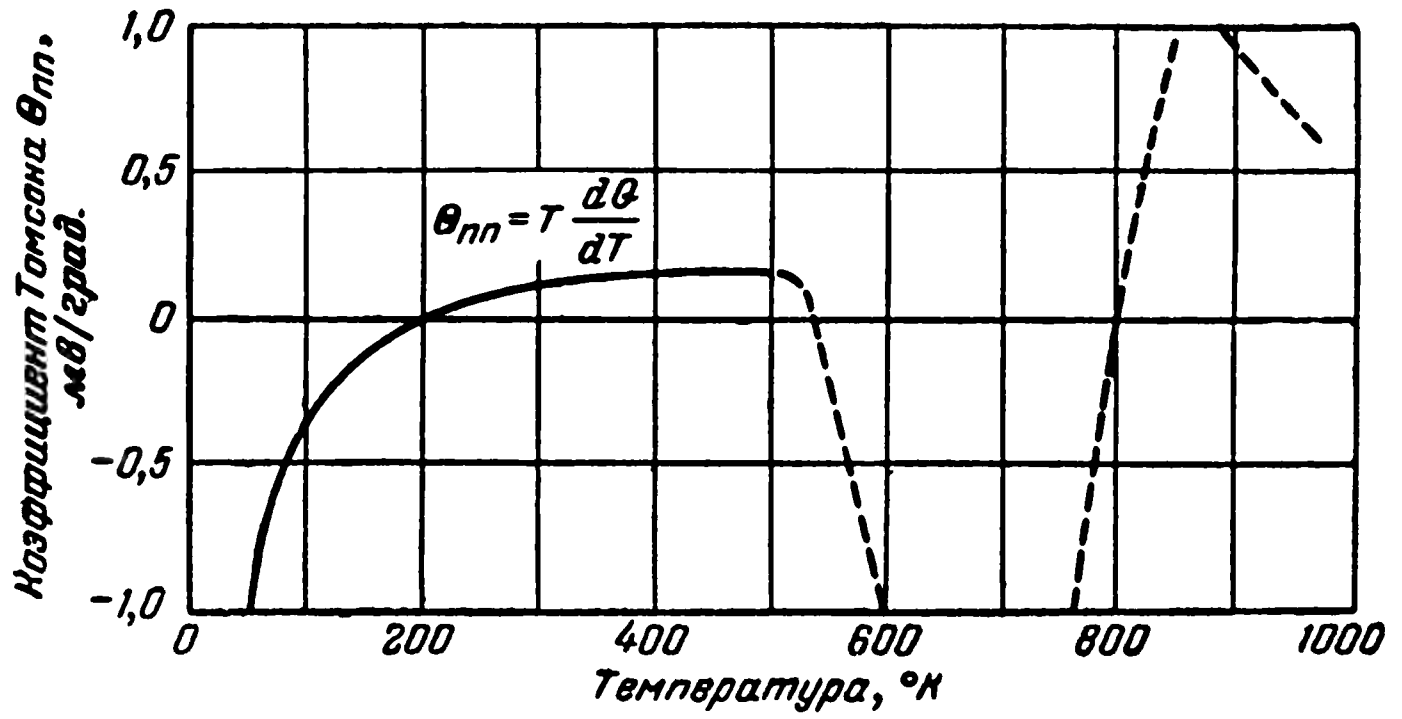


Рис. 23-7. Зависимость коэффициента Томсона от температуры, полученная из кривой Q на рис 23-5 на основании уравнения (23-17).

или акцепторов и энергию активации доноров и акцепторов. Одним словом, можно узнать довольно большое количество основных сведений о полупроводнике, получив и соответственным образом объяснив термоэлектрические зависимости. Если, кроме того, измерить зависимость проводимости от температуры, то мы получим необходимые данные для определения подвижности основных носителей.

23-5. СРАВНЕНИЕ С ЭКСПЕРИМЕНТАЛЬНЫМИ ДАННЫМИ

На рис. 23-8 приведены кривые зависимости термоэлектрического коэффициента Q от температуры в диапазоне примерно от — 150 до +650° С для трех образцов германия p-типа, активированного алюминием [Л. 2]. На этих кривых заметны все особенности теоретически предсказанной кривой, приведенной на рис. 23-5, за исключением поведения при температурах, лежащих ниже самых низких температур, достигнутых в эксперименте. Кривые приведены для трех различных концентраций активирующих примесей.

Даже поверхностное рассмотрение этих кривых позволяет легко выявить ряд явлений, согласующихся со сведениями о полупроводниках, полученными нами в этой и предыдущих главах. Во-первых, кривые, несомненно, переходят из области положительных в область отрицательных значений Q в согласии с ожиданием,

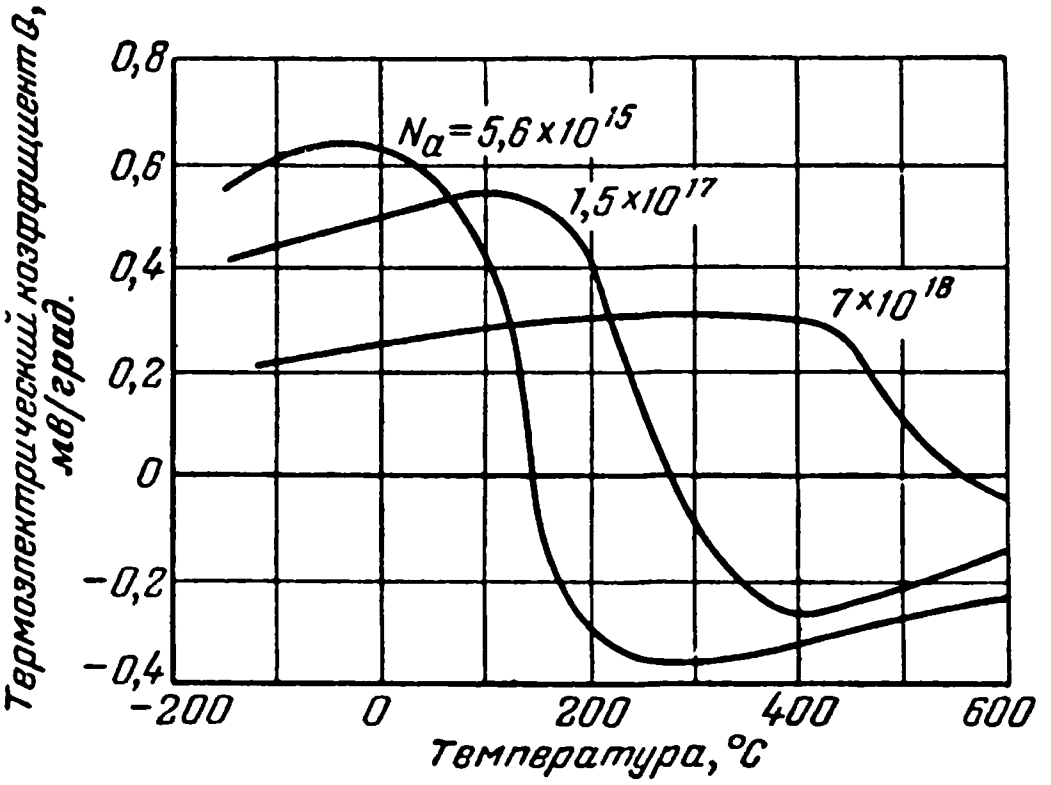


Рис 23-8 Экспериментальные кривые Q для германия p-типа; сравните их с формой теоретической кривой Q на рис. 23-5.

основанным на известном факте, что подвижность электронов в германии больше, чем подвижность дырок. Во-вторых, эта инверисия имеет место при более низких температурах для образцов с меньшим количеством примесей, что указывает на переход этих образцов в компенсированное состояние при более низких температурах в соответствии с нашими прежними сведениями об этом явлении, иллюстрированными на рис. 17-6. В-третьих, термоэлектрический коэффициент Q в области низких температур больше для образцов с меньшим содержанием примесей. Этого также следовало ожидать, так как у таких образцов уровень Ферми с ростом температуры поднимается быстрее, чем у образцов с бóльшим содержанием примесей.

ЗАДАЧИ

1. Покажите, как можно скомбинировать измерения удельного сопротивления и термоэлектрического коэффициента, чтобы получить независимое определение подвижности основных частиц в полупроводнике.

2. У металлов энергетические зоны вблизи уровня Ферми непрерывны. Почему же в контакте двух металлов возникают эффекты нагревания и охлаждения Пельтье?

3. Какой должна быть термо-э д с. в термоэлектрической цепи со стержнем из кремния *p*-типа, содержащим 10^{16} акцепторов в 1 $см^3$, причем его холодный контакт находится при температуре тающего льда, а горячий контакт имеет температуру 50° С? (Ответ: около 40 *мв*.)

4 Начертите качественные графики зависимости Q от T для двух или трех разных концентраций доноров в германии *n*-типа в диапазоне температур, включающем области, определяемые примесями, компенсированной и переходной. Учтите, что в германии подвижность электронов больше подвижности дырок. Сравните Ваши кривые с экспериментальными кривыми для германия *n*-типа, приведенными на рис. 23-8.

ЛИТЕРАТУРА

1. Starling S. G, Elements of Electricity and Magnetism, Longmans, Green, New York, 1931, chapter 13; Bridgman P. W., Thermodynamics of Electrical Phenomena in Metals, Macmillan, New York, 1934, chapter 2.

2. Middleton A. E, Scanlon W. W., Measurement of the Thermoelectric Power of Germanium at Temperatures Above 78° K, Physical Review, 1953, vol. 92, p. 219.

ГЛАВА ДВАДЦАТЬ ЧЕТВЕРТАЯ

ЭФФЕКТ ХОЛЛА

24-1. ПРОИСХОЖДЕНИЕ НАПРЯЖЕНИЯ ХОЛЛА

В § 2-7 кратко описан эффект Холла. Измерив направление напряжения Холла, можно определить, является ли полупроводник электронным (*n*-типа) или дырочным (*p*-типа). Кроме того, определив величину константы Холла R_H, можно получить другие полезные сведения относительно процессов в образце. Темой этой главы является более подробное рассмотрение этого вопроса. Первым шагом для этого должен быть вывод формулы для напряжения Холла, чтобы объяснить его происхождение и выявить физический смысл константы Холла R_H.

Предположим, что электрическая цепь сконструирована так, как показано на рис. 24-1, и содержит прямоугольный образец полупроводника *p*-типа, который помещен в однородное магнитное поле H_y, направленное нор-

мально к плоскости рисунка от нас. Если выключатель замкнуть, то в полупроводнике начинает течь слева направо ток положительных дырок I_x. В элементарном доказательстве, которое приведено ниже, мы будем прене-

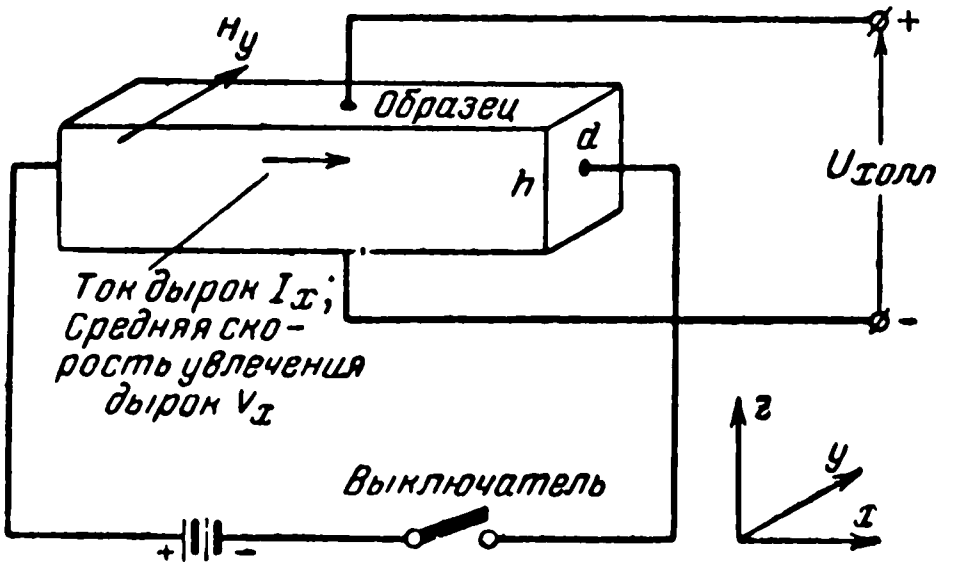


Рис. 24-1. К выводу уравнений, относящихся к эффекту Холла.

брегать беспорядочными тепловыми скоростями дырок и для простоты предположим, что каждая дырка движется в токе со средней скоростью смещения v_x. Заряженная частица e, движущаяся с этой скоростью в поперечном магнитном поле H_y, испытывает силу

$$F_x = \frac{e}{c}(v_x H_y),$$

где c — скорость света. Эта сила является той самой магнитодвижущей силой, которая приводит в действие стрелку амперметра и заставляет вращаться якорь электродвигателя. В обсуждаемом примере эта сила направлена в вертикальном направлении до z; вследствие этого траектории дырок изогнуты кверху, тогда как при отсутствии магнитного поля они в среднем были бы прямыми. Поэтому в очень короткий срок верхняя поверхность образца становится положительно заряженной по отношению к нижней поверхности из-за накопления дырок вверху вследствие этого искривления их траекторий. Накопление продолжается до тех пор, пока не установится достаточная разность потенциалов между верхней и нижней поверхностями, для того чтобы создать направленное вниз вертикальное электрическое поле E_z такой напряженности, которая могла бы уравновесить направленную вверх силу магнитного воздей-

ствия на положительные дырки. Когда устанавливается это стабильное состояние, то

$$\frac{e}{c}(v_x H_y) = eE_z = \frac{eV_{\text{Холла}}}{h}, \qquad (24\text{-}1)$$

где h — размер образца в направлении, перпендикулярном как току, так и магнитному полю.

В этом устойчивом состоянии средние траектории дырок снова делаются прямыми, идущими в направлении X. Если образец резко выраженного p-типа, так что можно пренебречь электронной составляющей тока, то можно написать:

$$I = pev_x dh, \qquad (24\text{-}2)$$

где h и d — размеры поперечного сечения образца.

Сравнивая уравнения (24-1) и (24-2), получаем:

$$V_{\text{Холла}} = \frac{I_x H_y}{cped}. \qquad (24\text{-}3)$$

Если обозначить через $R_H = 1/pe$ постоянную Холла для данного образца и помножить правую часть уравнения (24-3) на коэффициент 300, чтобы получить размерность в вольтах, то

$$V_{\text{Холла}} = \frac{10^{-8} R_H I_x H_y}{d}, \qquad (24\text{-}4)$$

где $R_H = \frac{1}{pe}$.

Если изображенный на рис. 24-1 полупроводник не p-, а n-типа, то основная часть тока создается электронами, движущимися справа налево. Тогда перед e и v_x в уравнении (24-1) следует поставить знак минус. Следовательно, траектории электронов будут и в этом случае искривляться кверху, как это имело место для дырок в образце p-типа; при этом верхняя поверхность будет накапливать отрицательный заряд вместо положительного. Следовательно, напряжение Холла будет отрицательным, а не положительным. Наблюдая направление напряжения Холла, можно, таким образом, сразу определить, к какому типу, n или p, относится полупроводник, если он имеет температуру, при которой состояние его определяется примесями. Во всем остальном вычисления,

приводящие к уравнению (24-4), в случае полупроводников n-типа те же, что и раньше; величина константы Холла оказывается равной:

$$R_H = -\frac{1}{ne}.$$

Приведенный выше элементарный вывод уравнения для напряжения Холла упрощает картину проводимости тока предположением, что v_x является постоянной скоростью смещения. Принимая во внимание, что v_x — лишь усредненная за длительный промежуток времени мгновенных величин скоростей смещения частиц, накладывающихся на случайные тепловые скорости, и усредняя скорость смещения для всех таких тепловых скоростей, получим дополнительный множитель $3\pi/8$ в выражении для R_H, и тогда имеем:

$$\left.\begin{aligned} R_H &= \frac{3\pi}{8pe} \text{ — для полупроводника } p\text{-типа} \\ \text{или} \\ R_H &= -\frac{3\pi}{8ne} \text{ — для полупроводника } n\text{-типа.} \end{aligned}\right\} \quad (24\text{-}5)$$

Из уравнения (24-5) видно, от чего зависит константа Холла R_H. Она различна для разных веществ и даже для разных образцов и для одного и того же вещества зависит от степени активирования вещества примесями n-типа или p-типа. R_H изменяется также в зависимости от температуры в состоянии, определяемом примесями, вследствие температурной зависимости p или n в этой области. Из величины константы Холла можно вычислить концентрацию электронов или дырок внутри полупроводника. Измеряя напряжения Холла в некоторой области температур, можно получить экспериментальные данные для построения графика зависимости концентрации носителей от температуры в этой области. Пользуясь этими данными, можно вычислить концентрацию и энергию активизации доноров и акцепторов.

Одновременно с измерением напряжения Холла в полупроводнике обычно стремятся измерить и электропроводимость образца. С помощью одновременно определенных константы Холла и проводимости можно найти не только концентрацию носителей, но и их подвижность.

Вычисление подвижности производится из выражений $\sigma = ep\mu_p$ или $\sigma = en\mu_n$, в которых p или n определены из измерений константы Холла. Таким образом, комбинируя константу Холла и проводимость, мы получаем:

$$\left.\begin{aligned} p \text{ или } n &= \pm \frac{3\pi}{8eR_H}; \\ \mu_p \text{ или } \mu_n &= \pm \frac{8R_H\sigma}{3\pi}, \end{aligned}\right\} \qquad (24\text{-}6)$$

где знак плюс относится к образцам p-типа, а минус — к образцам n-типа. Определенная таким образом подвижность иногда расходится с подвижностью, определенной для того же материала методом времени перемещения, описанного в § 18-3, пример 2А. Этим методом измеряется время перемещения инжектированных частиц на определенное расстояние в образце с известным ускоряющим полем. Оба значения подвижности, если они не совпадают, называются различно: «подвижность Холла» и «подвижность дрейфа». Считается, что в таких случаях частицы, подвижность дрейфа которых измерялась методом времени перемещения, значительную часть времени находятся в ловушках в кристаллической решетке. Согласно этому представлению носитель задерживается в такой ловушке на неопределенное время. По истечении этого времени он освобождается и снова присоединяется к потоку частиц, образующих электрический ток, совершая в нем несколько свободных пробегов и ряд соударений и перемещается на небольшое расстояние в направлении электрического поля. Рано или поздно эта частица снова захватывается другой ловушкой, затем снова освобождается и т. д. Таким образом, частичка свободно смещается в поле только в течение неопределенных интервалов времени между захватами. Подвижность, измеренная при таких обстоятельствах, конечно, меньше «истинной» подвижности частицы в свободном состоянии. С другой стороны, значение подвижности Холла дает подвижность частиц в свободном полете.

До сих пор мы делали выводы и объясняли измерения напряжения Холла, исходя из предположения, что имеем дело с образцами резко выраженного n-типа или p-типа, в которых концентрация неосновных носителей пренеобрежимо мала по сравнению с концентрацией

основных носителей. Поэтому вычисленные выражения пригодны только для состояния полупроводника, определяемого примесями, и для переходных состояний. Всякий раз, когда неосновные носители сильно влияют на движение частиц в полупроводнике при поперечном магнитном поле, так что необходимо принимать во внимание оба типа носителей, расчет константы Холла становится более сложным. В результате такого вычисления получается следующее выражение:

$$R_H = -\frac{3\pi}{8e} \cdot \frac{n(\mu_n/\mu_p)^2 - p}{[n(\mu_n/\mu_p) + p]^2}. \qquad (24\text{-}7)$$

В этом диапазоне температур надо, конечно, использовать также уравнение полной электропроводимости в виде:

$$\sigma = e(p\mu_p + n\mu_n).$$

Уравнение (24-7) приводит к уравнению (24-6) для образцов с хорошо выраженной активностью n-типа или p-типа в определяемых примесями состояниях. В кремнии и германии подвижность электронов превышает подвижность Холла. Поэтому если взять образец p-типа одного из этих материалов с большим количеством примесей и при низкой температуре, то получим положительную константу Холла. При повышении температуры такого образца до состояния компенсирования знак константы Холла становится отрицательным согласно уравнению (24-7) из-за большей подвижности электронов, хотя концентрация дырок остается несколько большей, чем концентрация электронов.

24-2. ИЗМЕРЕНИЕ ВЕЛИЧИН НАПРЯЖЕНИЯ И КОНСТАНТЫ ХОЛЛА

Схема расположения электродов для таких измерений показана на рис. 24-2. Напряжение Холла измеряется электродами, помещенными точно друг против друга на верхней и нижней поверхностях образца. Проводимость определяется по измерениям напряжения между зондами, расположенными у концов образца, при известном токе I_x через него. Образец с электродами помещается в среду с желаемой температурой. Спай термо-

пары приводится в контакт с поверхностью образца для измерения его температуры.

Измеренные так напряжения Холла лежат в пределах от нескольких микровольт до нескольких милливольт. Иногда они столь малы, что при их измерении возникают значительные трудности. Поэтому в современных установках для измерения эффекта Холла применяются устройства для возбуждения переменного магнитного поля в магнитной катушке; получаемое перемен-

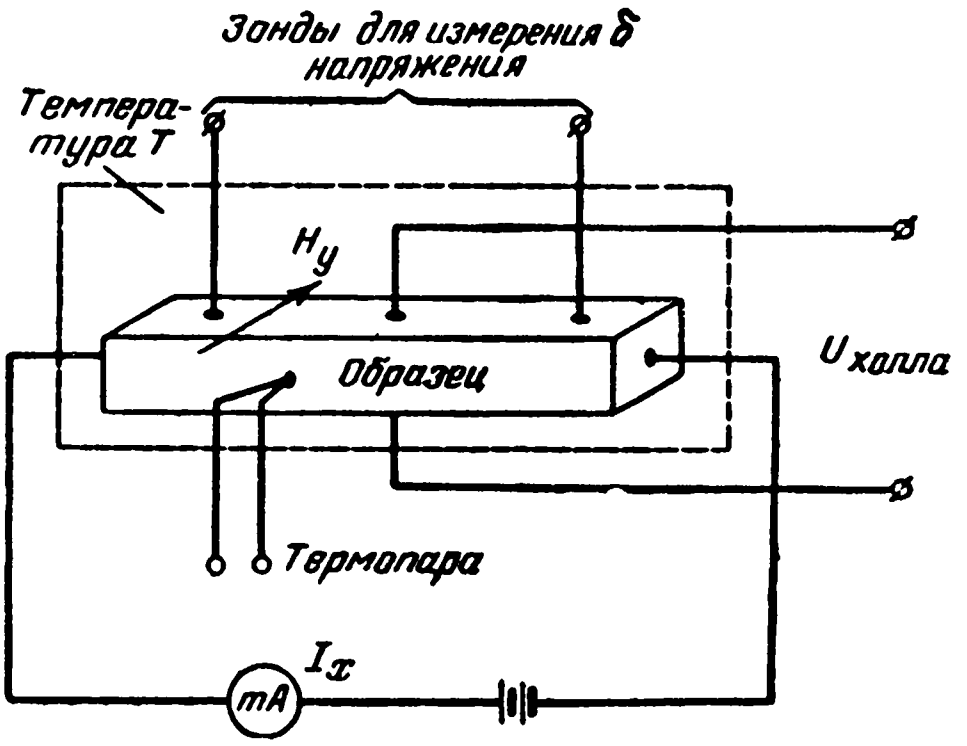


Рис 24-2. Схема экспериментальной установки для одновременного измерения R_H и σ на постоянном токе.

ное напряжение Холла снимается с контактов Холла на образце, усиливается, выпрямляется и, наконец, отсчитывается на калиброванном выходном приборе постоянного тока.

Измеряя константы Холла, следует учитывать некоторые важные практические соображения. Возможно, что самым важным из них после обеспечения необходимой точности измерения I_x, H, $U_{\text{Холла}}$ и размеров образца является расположение контактов для измерения напряжения Холла на образце в направлении X точно друг против друга. Если эта предосторожность не выполнена, то оба контакта будут находиться на разных эквипотенциальных плоскостях, когда по образцу течет ток I_x (даже при отсутствии магнитного поля), и между контактами Холла появится ложное напряжение. Так как очень трудно поместить эти контакты точно в нужном

месте, то ошибка, связанная с ложным напряжением, может оказаться равной, а иногда и больше напряжения Холла, измеряемого в присутствии магнитного поля. Если ложного напряжения нельзя избежать, то его надо тщательно измерить в отсутствие магнитного поля и вводить поправку на него при всех последующих измерениях При измерении напряжения Холла на постоянном токе можно избежать влияния этого ложного напряжения, определив напряжения Холла при наличии постоянного магнитного поля последовательно в прямом и обратном направлениях и затем усреднив оба результата. Влияние ложного напряжения может быть исключено также применением метода переменного магнитного поля, описанного в предыдущем параграфе.

Техника эксперимента должна также предусматривать возможность отличить напряжение Холла от всех других напряжений, возникающих между контактами Холла. Эти напряжения возникают в результате других «поперечных гальваномагнитных и термомагнитных явлений», которые хотя и мало известны, но тем не менее существенны [Л. 1] Одним из них является термоэлектрическая ошибка (эффект Эттингсхаузена). В опыте с эффектом Холла основные носители заряда в токе, проходящем через образец, не только создают ток, но и переносят тепло. Более медленные частицы имеют более изогнутые траектории в магнитном поле сравнительно с более быстрыми частицами. Поэтому на верхней поверхности образца (рис. 24-1) скопляются в основном медленные частицы и верхняя поверхность оказывается несколько холоднее нижней. Оба зонда для измерения напряжения Холла, находящиеся в контакте с этими поверхностями, имеют слегка различную температуру, вследствие чего появляется термо-э д с., добавляющаяся к истинному напряжению Холла. Другим источником погрешностей, который надо изучить и исключить, является то обстоятельство, что основной ток на концевых контактах вызывает нагревание Пельтье на одном конце и охлаждение Пельтье — на другом. Следовательно, оба конца имеют разные температуры и тепло течет по образцу от одного конца к другому. Поток тепла в направлении X в магнитном поле, имеющем направление Y, создает небольшое электрическое поле в направлении Z (эффект Риги-Ледука), что опять вызывает по-

явление нежелательной разности потенциалов между электродами Холла. С напряжением Риги-Ледука связано явление дополнительного поперечного температурного градиента, что еще более усложняет термоэлектрическую ошибку. Эти поперечные явления имеют сложную взаимосвязь, и поэтому трудно определить степень их влияния на измерения напряжения Холла [Л. 2]. В литературе [Л. 3] описан метод измерения напряжения Холла на переменном токе, не зависящий от других эффектов. В этом случае через образец пропускается переменный ток, питание магнита производится переменным напряжением другой частоты, а измерение напряжения Холла производится с помощью детектора с фильтром, пропускающим только суммарную или разностную частоту.

24-3. РЕЗУЛЬТАТЫ

Существует много работ [Л. 4—7] по использованию измерений эффекта Холла для исследования параметров полупроводниковых материалов. В качестве примера

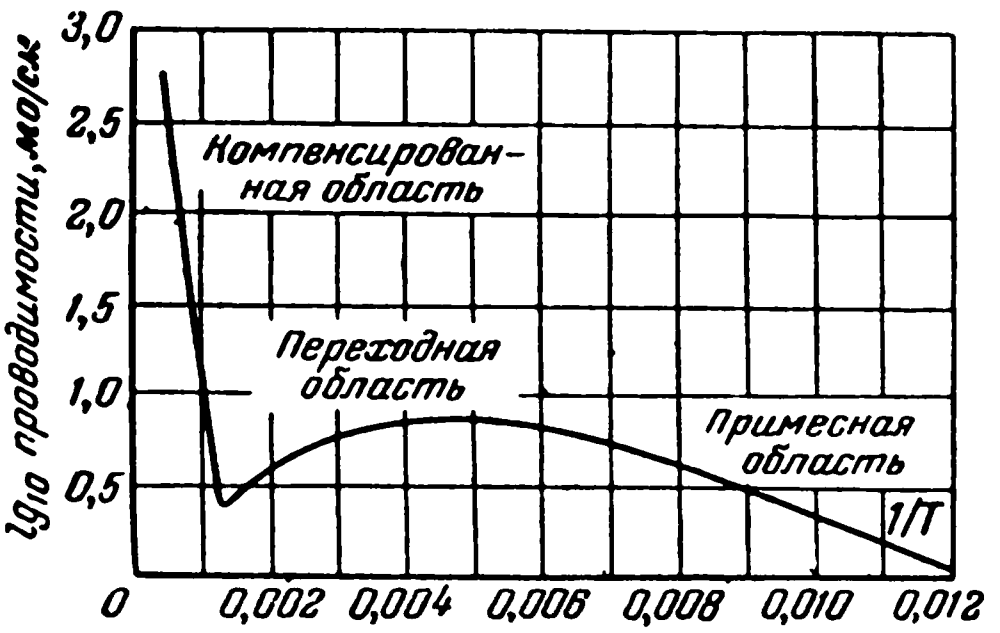


Рис 24-3 Кривая проводимости для образцов кремния p-типа с примесью бора, $N_a = 6{,}7 \cdot 10^{-17}$ атомов бора в 1 $см^3$.

интерпретации данных эффекта Холла на рис. 24-3, 24-4 24-5 и 24-6 приведены кривые, представляющие собой экспериментальные данные для образца кремния, активированного бором [Л. 5]. Рассмотрим, чго можно получить с помощью этих кривых. Испытанный образец был p-типа с концентрацией акцепторов (атомов бора) $6{,}7 \cdot 10^{17}$ $см^{-3}$. Эти атомы бора добавлялись во время изготовления образца. На рис. 24-3 показана проводи-

мость образца, а на рис. 24-4 — константа Холла, определенная экспериментально и нанесенная на график в функции обратной температуры $1/T$ в диапазоне приблизительно от 80 до 1 200° К. На кривой проводимости видны следующие характерные особенности: 1) асимптотическое приближение к прямой линии при низких температурах в состояниях, определяемых примесями; 2) сложное поведение в переходной области температур, в которой температурный коэффициент проводимости

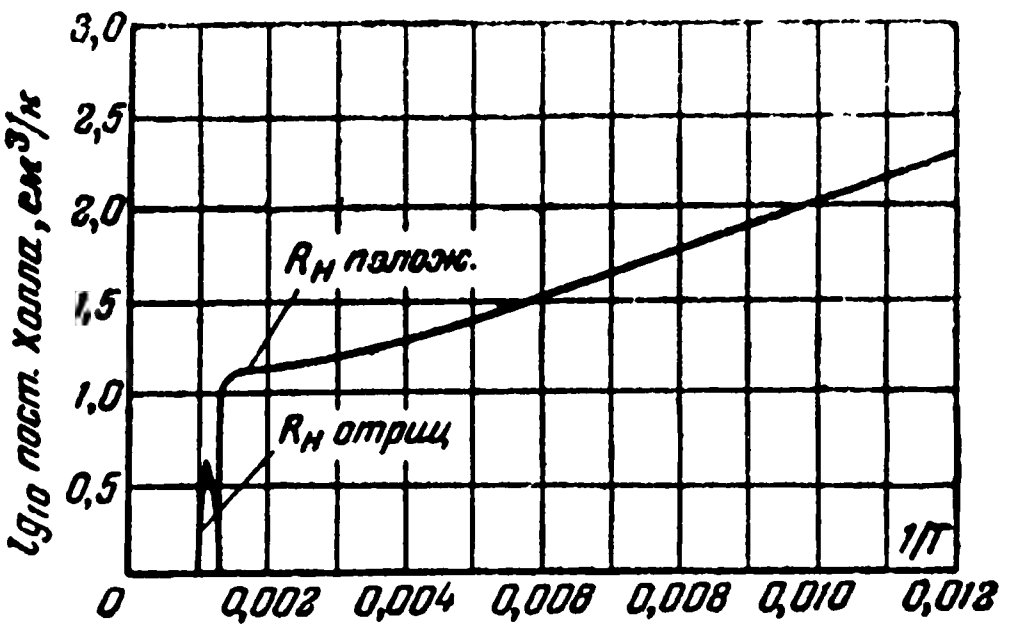


Рис. 24-4. Кривая постоянной Холла, полученная для того же образца кремния, для которого построена кривая проводимости на рис. 24-3 Значение R_H изменяется от положительного к отрицательному, когда образец становится компенсированным

меняет знак из-за зависимости подвижности частицы от температуры, и 3) прямолинейный отрезок, соответствующий компенсированному состоянию. Кривая константы Холла положительна при температурах ниже примерно 750° К. Выше этой температуры инверсии влияние электронов проводимости, число которых увеличивается вследствие теплового возбуждения из валентной зоны кристалла кремния, начинает преобладать над влиянием положительных дырок из-за большей их подвижности, и константа Холла приобретает результирующий отрицательный знак.

На рис. 24-5 приведена кривая зависимости концентрации носителей в функции обратной температуры. Эта кривая рассчитана по кривой константы Холла с помощью уравнения (24-6) для низких температур и урав-

нения (24-7) для высоких температур [1]. Из наклона этой кривой концентрации частиц при достаточно низких температурах в области, определяемой влиянием примесей, можно вычислить энергию активации акцепторов E_a. Это вычисление можно выполнить с помощью уравнения (17-12). Для данного образца кремния $E_a = 0{,}05$ *эв*. В переходной области температур, где акцепторы заполнены

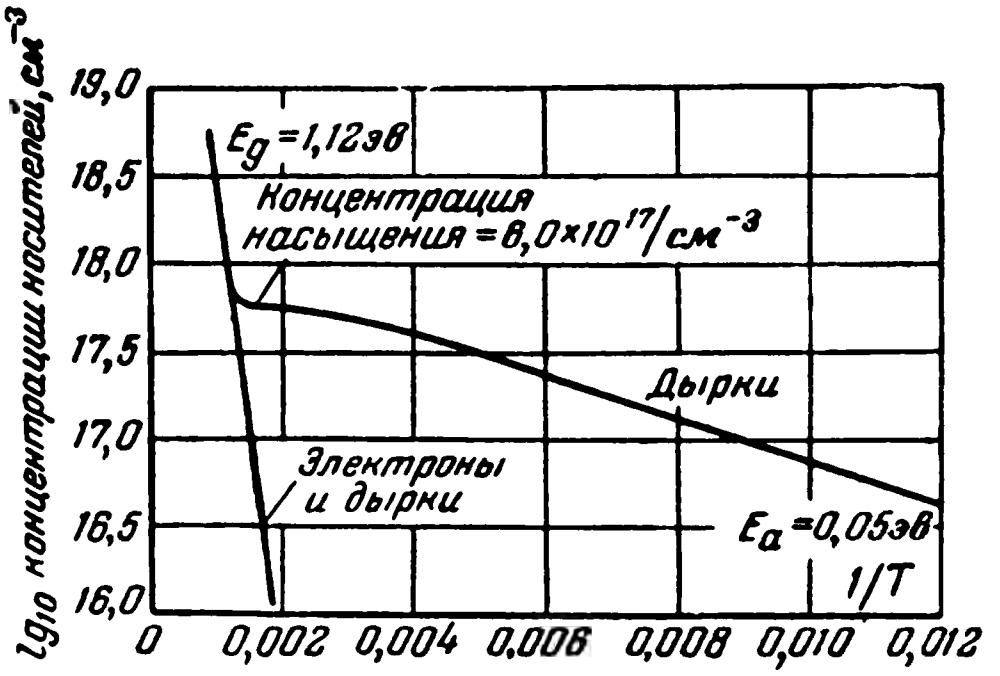


Рис 24-5. Зависимость концентрации носителей, вычисленная из данных на рис. 24-4.

до состояния насыщения, концентрация положительных дырок равна, как видно, $6{,}0 \cdot 10^{17}$ $см^{-3}$. Эта величина очень близка к концентрации примешанных акцепторов (атомов бора) ($6{,}7 \cdot 10^{17}$ $см^{-3}$), и можно сказать, что каждый атом бора в действительности участвует одной положительной дыркой в механизме проводимости в этой области температур. Из наклона кривой в области компенсированного состояния можно вычислить ширину E_g запрещенной зоны кремния. Согласно уравнению (17-7) это вычисление дает $E_g = 1{,}12$ *эв*.

На рис. 24-6 показана кривая зависимости подвижности от обратной температуры для этого образца. Эта кривая подвижности приближается и совпадает с кривой $\mu_p \sim 1/T^{3/2}$ при температуре гораздо ниже той, при

---

[1] При высоких температурах, когда приходится учитывать также и электроны проводимости, надо вычислять четыре величины (n, p, μ_n и μ_p) из измерений только двух величин (σ и R_H). Такие вычисления кажутся невозможными, однако при достаточно высоких температурах $n \approx p$, и если предположить, что кривая μ_p в этой области температур является продолжением кривой μ_p при низких температурах и изменяется, как $1/T^{3/2}$, то число подлежащих вычислению независимых величин уменьшается до двух.

которой концентрация неосновных носителей становится достаточно большой, чтобы можно было использовать уравнение (24-7). Этот факт рассматривается как подтверждение того, что для компенсированного состояния справедливо соотношение $1/T^{3/2}$ и что определенными отсюда значениями подвижности можно пользоваться при вычислении концентрации носителей в этой области температур.

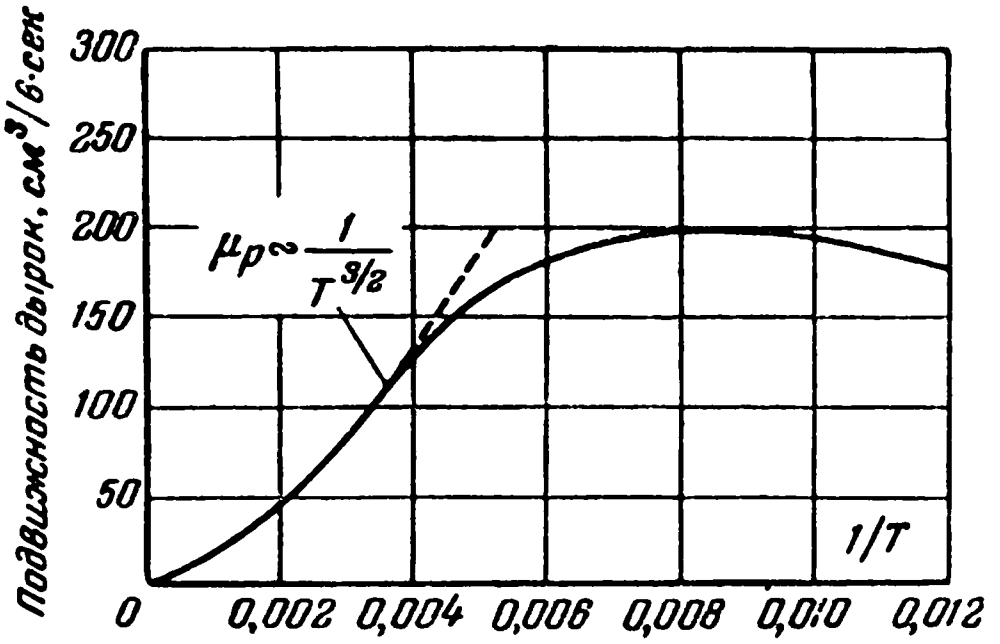


Рис 24-6 Подвижность дырок в образце кремния, вычисленная по данным измерений коэффициента Холла и проводимости, приведенным на рис 24-3 и 24-4.

Таким образом, измерения эффекта Холла представляют собой важный метод изучения свойств материалов. Сведения, получаемые из данных по эффекту Холла в комбинации с данными по исследованию проводимости, времени жизни, фотоэлектрических и термоэлектрических явлений, а также других наблюдений, могут быть использованы для построения чрезвычайно ясной, полной и согласующейся количественно картины изучаемого материала, особенно тех параметров, которые важны для разработки полупроводниковых приборов. Существенно отметить, однако, что развитие этих различных методов зависит от использования наших знаний об основных физических процессах в твердых телах. Вообразите, как бесплодны были бы наши попытки интерпретировать измерения, сделанные на полупроводниках, если бы мы не руководствовались современными представлениями о зонной структуре твердых тел, о роли примесей, о природе и поведении положительных дырок или о том, как на эти явления влияют тепловые, оптические, электрические и магнитные эффекты.

ЗАДАЧИ

1. Как бы Вы рассчитали положение уровня Ферми в примесной или переходной области температур для полупроводника n-типа, зная экспериментально определенное значение R_H? Что надо предусмотреть и какие дополнительные сведения Вам потребуются, чтобы произвести такие расчеты для области компенсации?

2. Покажите, как бы Вы вычислили константу Холла, зная термо-э. д. с. полупроводника.

3. Докажите, что в сильно активированном полупроводнике p-типа, по которому проходит ток в направлении, нормальном к магнитному полю, угол между током и эквипотенциальными плоскостями в полупроводнике равен $\operatorname{arcctg}(10^{-8}\mu_p H)$.

ЛИТЕРАТУРА

1. Bridgman P. W., The Thermodynamics of Electrical Phenomena in Metals, Macmillan, New York, 1934, chapter 7.

2. Lindberg Olof, Hall Effect, Proceedings of the Institute of Radio Engineers, 1952, vol. 40, p. 1414.

3. Russel B. R., Wahlig C., A New Method for the Measurement of Hall Coefficients, Review of Scientific Instruments, 1950, vol. 21, p. 1028.

4. Engelhard E., Electrical and Optical Properties of Semiconductors, Annalen der Physik, vol. 17, 1933, p. 501.

5. Pearson G. L., Bardeen J., Electrical Properties of Pure Silicon and Silicon Alloys Containing Boron and Phosphorus, Physical Review, 1949, vol. 75, p. 865.

6. Dunlap W. C., Some Properties of High Resistivity p-Type Germanium, Physical Review, 1950, vol. 79, p. 286.

7. Middleton A. E., Scanlon W. W., Measurement of the Thermoelectric Power of Germanium at Temperatures above 78° K, Physical Review, 1953, vol. 92, p. 219.

ГЛАВА ДВАДЦАТЬ ПЯТАЯ

НЕКОТОРЫЕ ВОПРОСЫ ТЕХНОЛОГИИ ИЗГОТОВЛЕНИЯ ПРИБОРОВ

Читатель теперь понимает, что сделать плоскостной транзистор n-p-n типа — достаточно сложная задача. Нельзя просто соединить полупроводники соответствующей формы n-типа и p-типа, чтобы получить триод. До сих пор никому не удалось сделать хороший плоскостной транзистор таким способом. Этот метод слишком груб для производства таких приборов, и задачи, которые необходимо разрешить для осуществления подобной технологии, чрезвычайно трудны. Соприкасающиеся поверхности должны быть точно параллельными и гладкими

с точностью до атомных размеров, чтобы обеспечить плотный контакт достаточно большой части соприкасающихся поверхностей. В процессе изготовления таких поверхностей нельзя нарушать химический состав или кристаллическое строение атомных слоев, находящихся непосредственно у этих поверхностей. Если требуется, чтобы переход был воспроизводимым и имел заранее предусмотренные свойства, с поверхностей должны быть удалены всякие окислы и другие химические пленки и они должны сохраняться в таком виде до тех пор, пока не будет осуществлен контакт.

Трудность выполнения всех этих требований привела к развитию специальных методов создания переходов в монокристаллах в процессе их выращивания, а также к образованию переходов внутри монокристаллов после завершения их выращивания. Методы создания таких переходов часто упоминались в предыдущих главах, и читатель уже в общих чертах знаком с основной разницей между понятиями выращивание кристалла, сплавление и диффузия. Однако в предыдущих главах об этом упоминалось только с такими подробностями, которые были необходимы, чтобы получить понятие об этом различии и о свойствах и пределыых возможностях изготовляемых приборов.

Задача этой главы — более подробное обсуждение этих процессов и освещение тех вопросов физики, химии и металлургии, от которых они зависят.

25-1. ПРИГОТОВЛЕНИЕ И ОЧИСТКА ГЕРМАНИЯ

Основным веществом для производства германия для полупроводников является дувокись германия GeO_2. Это вещество является одним из побочных продуктов при рафинировании цинка. В готовом виде GeO_2 с торговой точки зрения уже совсем чист; концентрация примеси в нем не превышает 0,001%. Германий получается восстановлением из двуокиси при нагревании в потоке водорода при температуре приблизительно 650° С. Во время этого процесса восстановления двуокись находится в графитовом тигле, который в свою очередь помещается в кварцевую трубку, через которую пропускается водород. Последней стадией «печного цикла» является прогрев до температуры 1 000° С для сплавления губчатой массы германия, получаемой после восстановления. Вы-

нутый из печи затвердевший германий имеет вид компактного поликристаллического слитка. Большинство металлических примесей, первоначально присутствовавших в двуокиси, сохраняется в германии, тогда как значительная часть таких примесей, как мышьяк и сурьма, удаляются во время восстановления в виде летучих гидридов.

На этой стадии германий коммерчески и промышленно совершенно чист, но все еще чрезвычайно загрязнен, если его рассматривать как исходный материал для изготовления полупроводниковых приборов. Он еще должен пройти процесс дальнейшей очистки, состоящий в повторных плавках и рекристаллизациях. Этот процесс эволюционировал в хорошо известный в настоящее время простой и эффективный метод зонной плавки [Л. 1].

Если порция германия плавится в длинном горизонтальном лотке, а затем затвердевает таким образом, что охлаждение начинается с одного конца лотка и медленно и постепенно распространяется к другому концу, то бóльшая часть остаточных примесей перемещается перед границей твердой и жидкой фаз и в последнем застывающем участке концентрация примесей оказывается выше средней. Затвердевание должно происходить медленно, чтобы примеси, которые скапливаются перед границей раздела твердой и жидкой фаз, успевали диффундировать в основную массу жидкой фазы; концентрация оставшихся примесей должна быть в последней равномерна, в противном случае эффективность рекристализации не будет оптимальной. Практически быстрота продвижения границы затвердевания германия должна быть порядка 0,025 *мм/сек* или еще медленнее в зависимости от геометрии лотка, его загрузки и интенсивности конвекционного или искусственного перемешивания жидкой фазы. Это отделение примесей является следствием стремления большинства атомов примесей сохраняться в жидком растворе, а не захватываться затвердевающей фазой.

Если конец рекристаллизованного слитка (который теперь содержит бóльшую часть примесей) отрезать, то оставшаяся часть содержит среднюю концентрацию примесей, меньшую той, которая была в первоначальном материале. Материал, очищенный таким образом, может стать еще более чистым, если повторить цикл плавления

и рекристаллизации. После шестикратного повторения этого цикла, причем количество материала уменьшается, вследствие того что каждый раз отрезается конец слитка, затвердевающий последним, германий достигает такой степени чистоты, что дальнейшее повторение процесса становится невыгодным, так как при каждом нагреве материал несколько загрязняется от лотка, стенок

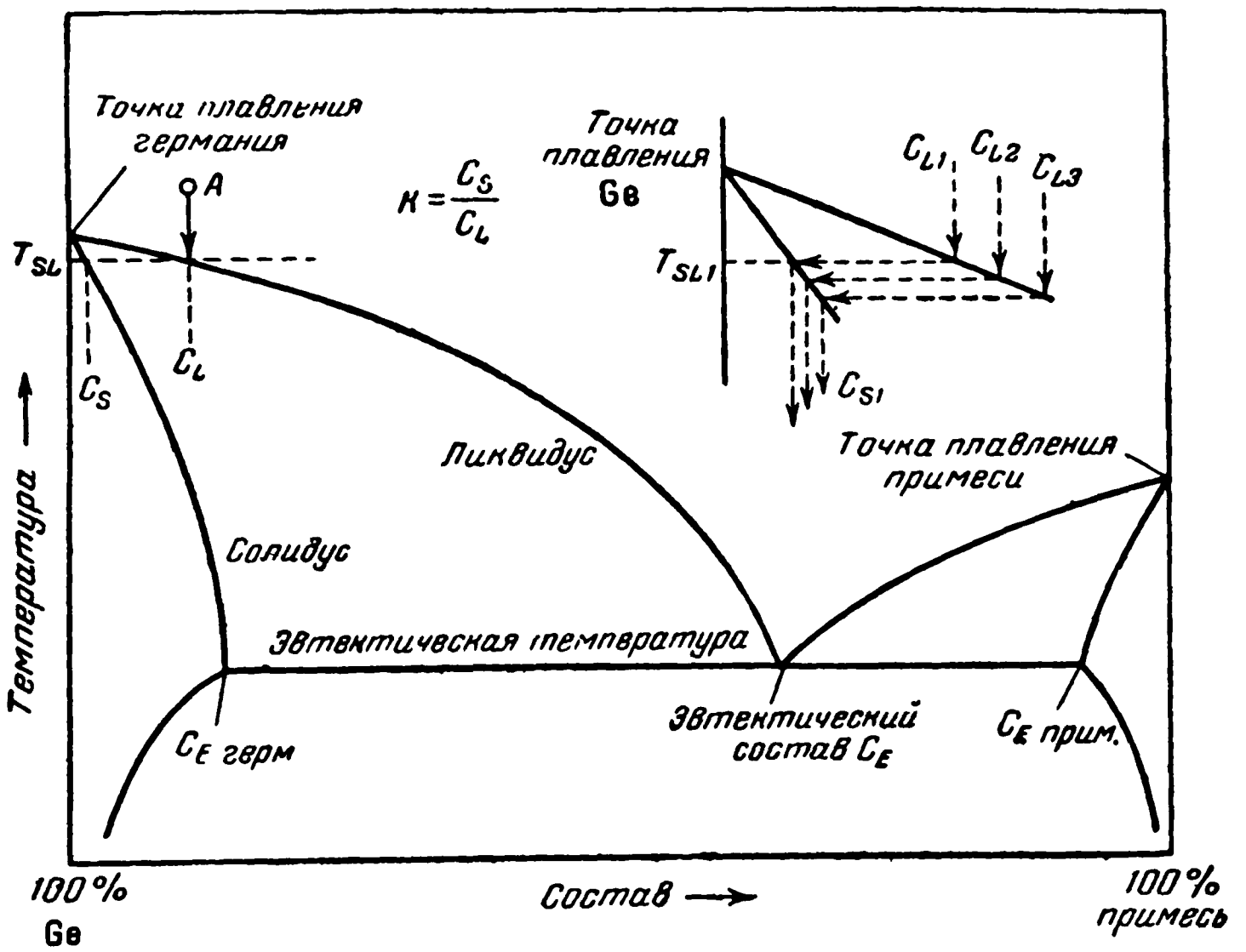


Рис. 25-1. Фазовая диаграмма дает полезную информацию о переходах от твердого к жидким состояниям для смеси двух (или более) веществ.

печи и «нейтральной» газовой атмосферы: Когда степень очистки достигла этой стадии, ее следует прекратить.

Чтобы глубже понять физико-химическую природу этого процесса разделения, лучше всего изучить фазовую диаграмму двухкомпонентной системы, приведенную в несколько идеализированном виде на рис. 25-1. Фазовая диаграмма является удобным графическим способом изображения всей суммы сведений о поведении твердой и жидкой фаз смеси двух или более веществ при различных температурах. Построение и изучение фазовых диаграмм само по себе очень интересный предмет, который мы, однако, не будем подробно разбирать. Опишем до-

статочно подробно лишь ту часть диаграммы на рис. 25-1, которая относится к процессу очистки.

Абсцисса фазовой диаграммы изображает состав рассматриваемой системы (в данном случае — системы германия и типичной примеси). Хотя неочищенный образец германия всегда содержит несколько видов примесей, каждая из которых проявляет себя особым образом, мы предположим сейчас, что имеется только один вид примеси.

По оси ординат фазовой диаграммы отложены температуры.

Чтобы видеть, как осуществляется очистка методом плавления и рекристаллизации, предположим, что мы начали с образца германия, содержащего небольшое количество примесей, обозначенного на диаграмме C_L, при температуре достаточно высокой, чтобы смесь была совершенно расплавлена, как показано в точке A. По мере понижения температуры точка A движется вниз, пока ее путь пересечется с кривой ликвидуса. При этой температуре T_{SL} смесь [вещество] начинает твердеть; но та часть, которая твердеет первой, имеет состав не C_L, а C_S, соответствующий пересечению ординаты температуры T_{SL} с кривой солидуса. Таким образом, твердое вещество чище, чем жидкое, из которого оно кристаллизуется. С другой стороны, оставшееся жидкое вещество становится теперь более загрязненным, и по мере того как продолжается затвердение, жидкая фаза последовательно проходит через стадии с содержанием C_{L1}, C_{L2} и т. д. (см. правый верхний угол на рис. 25-1). В то же время и соответствующие порции отвердевающего вещества становятся все более загрязненными: C_{S1}, C_{S2} и т. д., но в меньшей степени. Когда затвердевание при постепенно понижающейся температуре пройдет достаточно далеко, состав оставшейся жидкой фазы делается эвтектическим C_E. Затем это жидкое вещество отвердевает при постоянной эвтектической температуре уже без изменения состава до тех пор, пока весь расплав не затвердеет.

Твердое вещество, получившееся в результате отвердевания жидкости эвтектического состава, называется эвтектической смесью. В нашем случае она состоит из физической смеси кристаллов твердого раствора примеси в германии, соответствующего составу $C_{E,\,Ge}$, и кристал-

лов твердого раствора германия в примеси, соответствующего составу $C_{E,1\text{тр}}$.

С точки зрения процесса очистки германия нас интересует лишь часть этой фазовой диаграммы — ее самый левый конец, соответствующий концентрациям примеси, составляющей малые доли процента. В этой области кривые солидуса и ликвидуса являются практи-

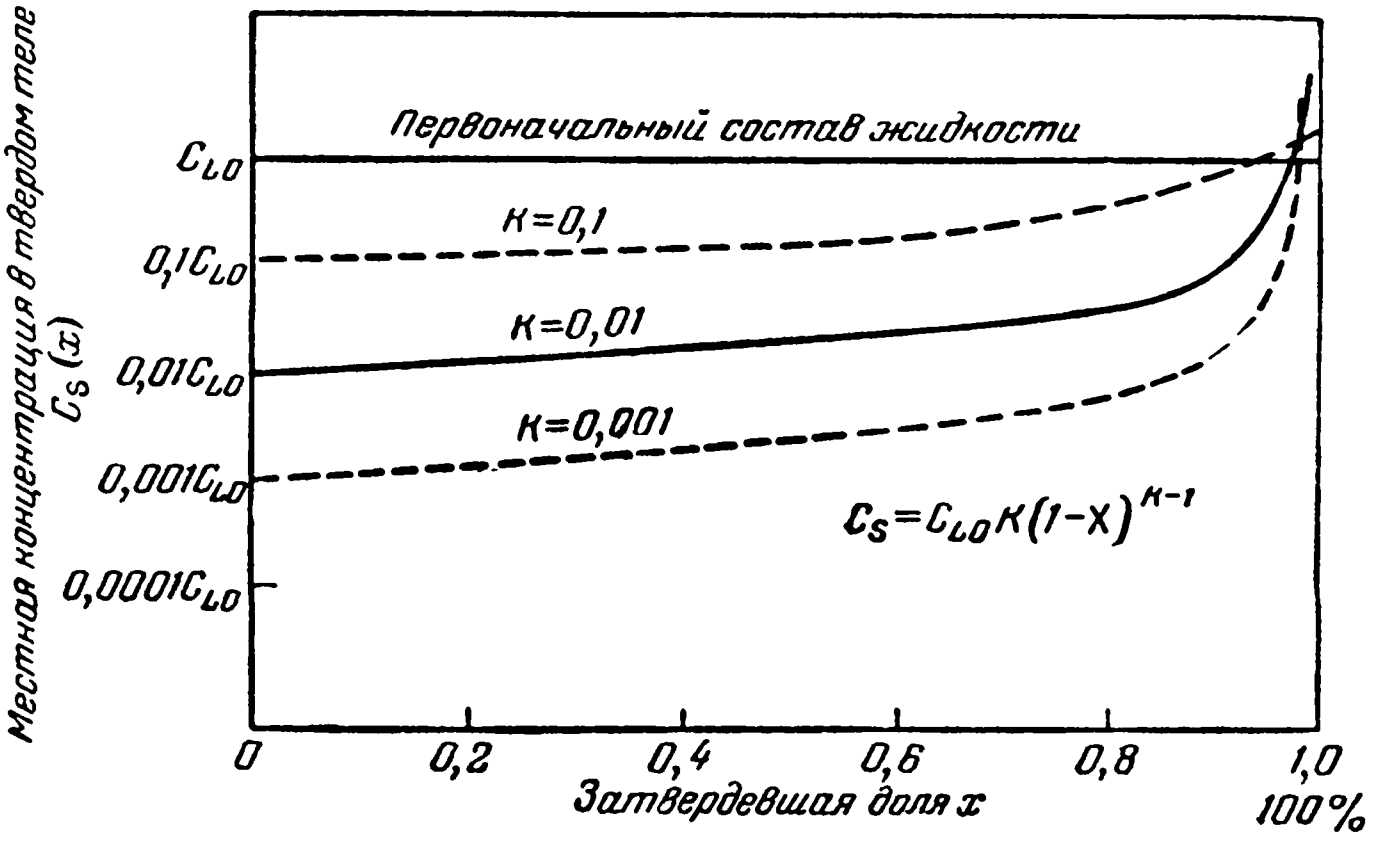


Рис. 25-2. Во время затвердевания загрязненного материала концентрация примеси в той части, которая должна затвердеть, зависит от относительной длины уже затвердевшей части слитка.

Примеси	K	Примеси	K
P	0,12	Sb	0,005
Ga	0,10	In	0,001
As	0,06	Ge	1,0

чески прямыми линиями, и мы можем определить для каждой примеси величину, называемую константой разделения K, как отношение C_S/C_L. Эта константа разделения дает количественную меру очистки, которой можно достичь для данной примеси за один цикл плавления и рекристаллизации. Сплошная кривая на рис. 25-2 изображает состав кристаллизовавшейся твердой фазы как функцию доли вещества, перешедшего в твердое состояние, для частного случая $K=0{,}01$. Пунктирные кривые изображают подобные же зависимости для других величин K. Все кривые вычислены с помощью уравнения [Л. 2 и 3]

$$C_S = C_{L0}K(1-X)^{K-1}, \qquad (25\text{-}1)$$

где C_{L0} — первоначальная концентрация примесей в расплаве, а X — доля уже затвердевшего вещества.

Это уравнение справедливо в предположении, что отвердевание происходит в условиях, обеспечивающих полную однородность распределения примеси в остающейся жидкой фазе.

Ясно, что очистка более совершенна и охватывает бóльшую часть рекристаллизованного слитка для примесей с меньшими константами K. Постоянные разделения K для некоторых примесей, имеющих значение при обработке германия, приведены на рисунке. Если в образце германия присутствует одновременно ряд различных примесей, каждая примесь будет изгоняться из образца с быстротой, соответствующей величине ее константы K. Если начать очистку образца германия с концентрациями примесей, подлежащих удалению, по 0,001% каждая, то достаточно одной или двух перегонок, чтобы полностью, кроме несущественных следов, удалить такие примеси, как индий и сурьма. Однако понадобится еще несколько циклов очистки, чтобы уменьшить содержание галлия и фосфора до безопасного уровня.

Физические свойства, от которых зависят величины K этих различных примесей, связаны частично с относительными размерами их атомов, в частично — с ориентацией их валентных орбит. Если атомы некоторой примеси больше атомов германия, то ясно, что вероятность включиться в растущий кристалл для таких атомов мала, так как это вызвало бы значительные деформации основной решетки. Такие примеси должны были бы иметь соответственно малые значения K. Если бы даже посторонний атом имел размеры, позволяющие ему легко встроиться в решетку германия, то могло бы оказаться еще, что направления его валентных связей не совпадают с соответствующими направлениями для соседних атомов германия без значительной деформации. Растущая граница между твердой и жидкой фазами стремилась бы столкнуть такие атомы, и эти примеси должны были бы иметь низкие значения K.

Процесс зонной очистки очень оригинально объединяет в одну операцию несколько ступеней плавления и рекристаллизации, необходимые для очистки германия. Кроме того, при этом процессе нет необходимости отре-

зать и выбрасывать конец слитка после каждого цикла В процессе зонной очистки слиток германия, находящийся в длинной лодочке, медленно продвигают через ряд индукторов. Внутри каждого индуктора происходит местное расплавление германия лишь на небольшом участке (зоне) по оси слитка. По мере того как слиток продвигается через каждый индуктор, соответствующая зона плавления как бы проходит вдоль всего слитка и очищает его в отношении C_S/C_L. На рис. 25-3 приведена схема установки зонной очистки.

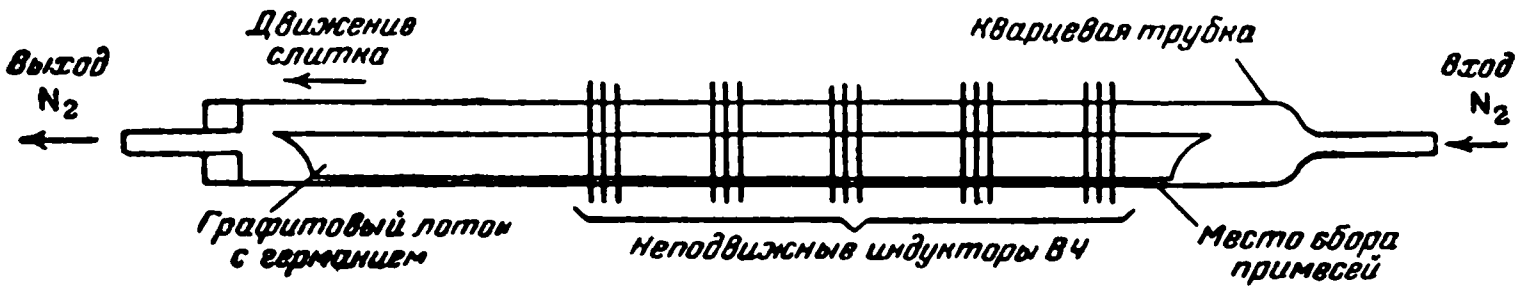


Рис. 25-3. Схема зоноочистительной установки, обеспечивающей медленное прохождение лотка с германием через ряд зонных индукторов, по которым протекает ток высокой частоты.

На рис. 25-4 показаны расчетные распределения концентрации примеси вдоль слитка, прошедшего процесс зонной очистки (в предположении, что поперечное сечение слитка постоянно). Предполагается, что имеется лишь один вид примеси с начальной концентрацией C_{L0}, для которой $K=0{,}1$, и что длина расплавленной зоны составляет 0,1 длины слитка. Верхняя жирная кривая дает распределение примеси после прохождения одной зоны, а остальные кривые — после 2-го, 3-го, 4-го и 5-го проходов зон соответственно. Следует отметить, что верхняя кривая поднимается несколько быстрее и круче, чем кривая для $K=0{,}1$ на рис. 25-2. Этого можно было ожидать, поскольку расплавленная зона, ограниченная по длине, быстрее загружается выбрасываемой примесью, чем это бы имело место в незатвердевшей части охлаждаемого обычным образом расплава, и поэтому оставляет за собой твердую фазу с быстрее возрастающей концентрацией примеси.

Если в приведенном выше примере $C_{L0}\approx 0{,}001\,\%$, что более или менее типично для концентрации галлия в первоначальном материале, то продвижение лодочки через пять индукторов даст в результате очищенный слиток с остаточной концентрацией галлия, пониженной до ме-

нее чем 1/100 000 000 на длине 0,4 полной длины слитка.

В производственной практике слитки, проходящие процесс зонной очистки, сперва полностью расплавляются, а затем медленно кристаллизуются, начиная с того конца, который первым вступает в индукторы. Эта пред-

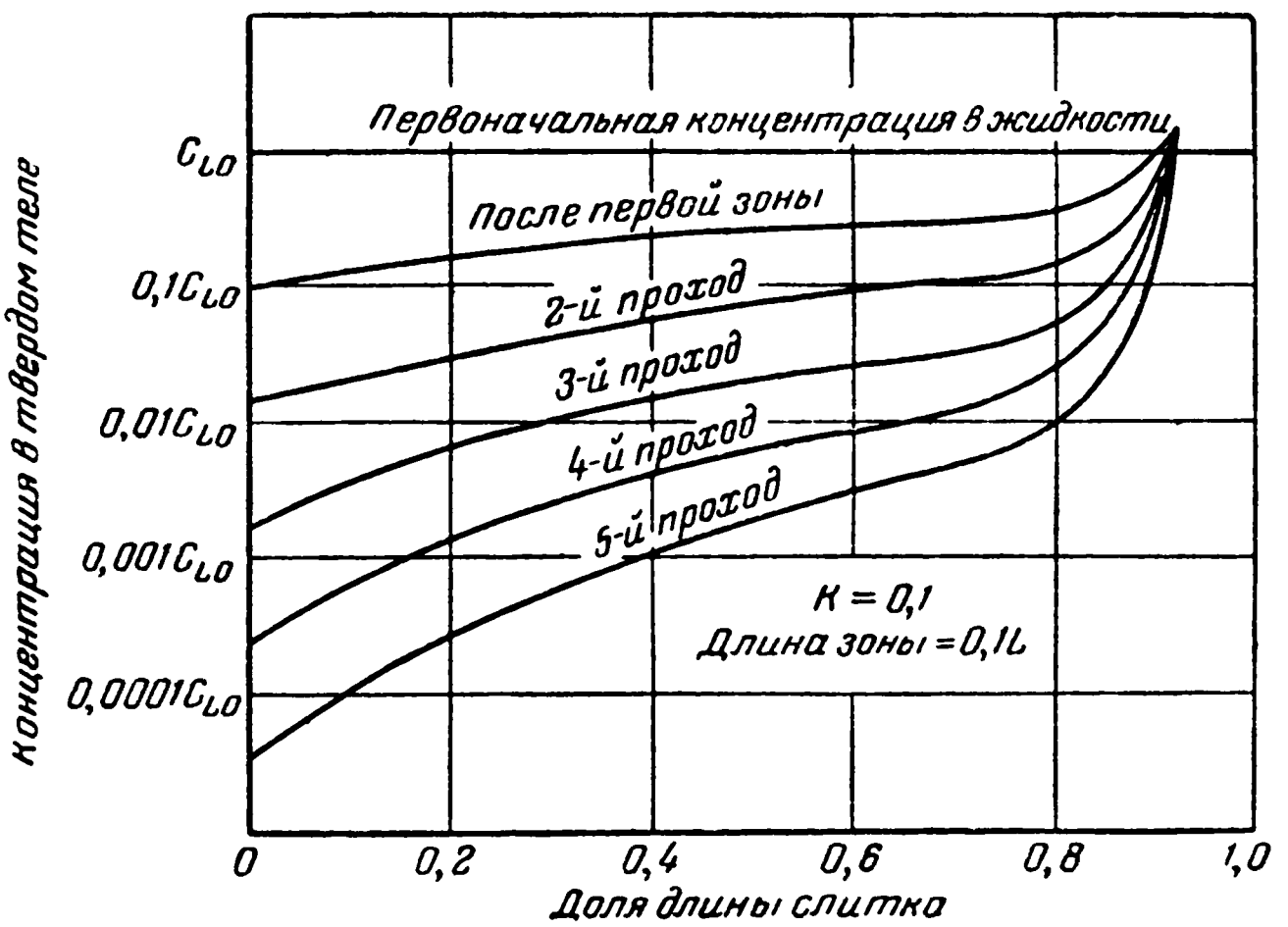


Рис. 25-4. Распределение примесей в слитке, прошедшем ряд последовательных зонных очисток, зависит от коэффициента разделения примеси, длины зоны, числа проходов через зоны и расстояния от начала очищаемого стержня.

варительная плавка дает в результате слиток, полностью совпадающий по форме с лодочкой, в которой он находится, что позволяет проводить последующий процесс зоной плавки без взбалтывания, встряхивания и других механических возмущений тех частей слитка, которые отделены жидкими зонами. Эта начальная кристаллизация является одновременно и первым циклом процесса очистки.

25-2. СОЗДАНИЕ ПЕРЕХОДОВ МЕТОДОМ ВЫТЯГИВАНИЯ КРИСТАЛЛОВ

Первым пригодным для практики и контролируемым методом создания *n-p-n* структур явился метод вытягивания кристаллов [Л. 4, 5 и 6], применяемый до сих пор при изготовлении низкочастотных и среднечастотных

транзисторов. Метод заключается в том, что в химически инертном тигле в атмосфере благородного газа расплавляется от 25 до 50 $см^3$ германия n-типа. В это расплавленное вещество опускается маленькое «зерно» монокристалла германия и затем медленно вытягивается кверху. По мере того как это зерно поднимается, германий на нем затвердевает; зерно, таким образом, растет и по-

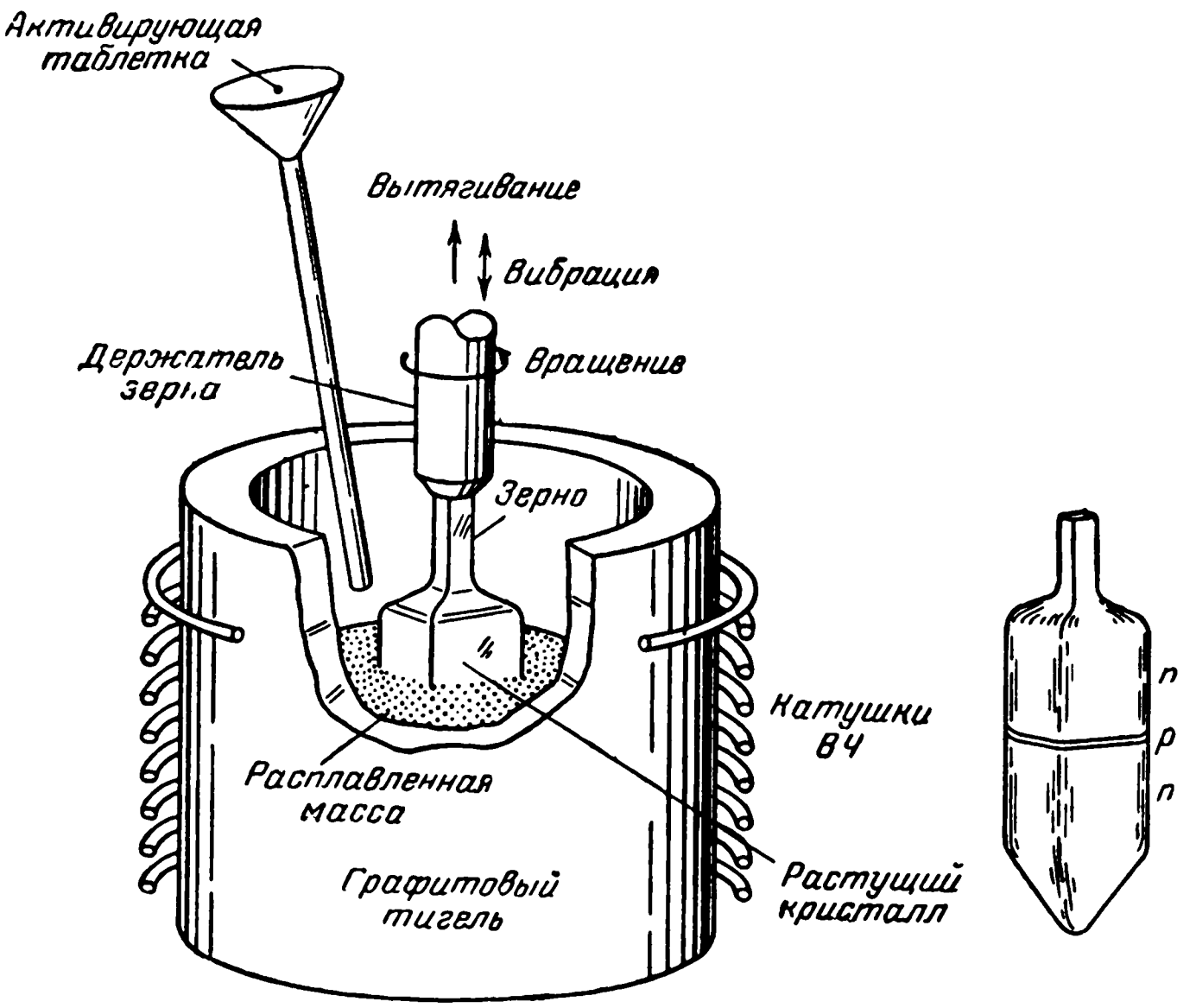


Рис. 25-5. Этот кристалл с n-p-n структурой получен методом вытягивания.

степенно поглощает все расплавленное вещество, если процесс не прекращается извне.

На рис. 25-5 приведена схема прибора для вытягивания кристалла. Форма вытягиваемого кристалла определяется в основном температурным градиентом на границе между твердой и жидкой фазами и скоростью, с которой вытягивается кристалл, причем обе эти величины должны быть надлежащим образом согласованы друг с другом и с температурой плавления вещества. На ранних стадиях выращивания скорость вытягивания замедляется до тех пор, пока диаметр кристалла не вырастет от диаметра зерна (примерно 2 *мм*) до требуемой

величины конечного кристалла, обычно порядка 18 *мм*. Затем вертикальный рост продолжается с такой скоростью, чтобы сохранить постоянный диаметр растущего кристалла в течение всего остального времени вытягивания. В настоящее время применяется скорость вытягивания порядка 0,02—0,05 *мм/сек*. Изготовленный кристалл имеет размер и вид небольшой короткой морковки, имеющей близкое к прямоугольной форме поперечное сечение, так как характерная для германия кубическая кристаллическая структура предопределяет различные скорости роста в разных кристаллографических направлениях. В большинстве случаев весь образец оказывается монокристаллом.

Если надо создать посредине такого кристалла во время его роста *p-n* переход, то в расплавленное вещество опускается небольшая таблетка из акцепторного материала. Количество его должно быть таким, чтобы подавить порвоначальный электронный характер расплава и создать в нем желаемую избыточную концентрацию акцепторов. С этого момента образец становится кристаллом *p*-типа. Если надо получить триод с *n-p-n* структурой, то через 1—2 *сек* в расплавленную массу опускается еще одна таблетка, содержащая донорный материал в количестве, достаточном для того, чтобы в расплавленной массе создался желаемый излишек доноров, и кристалл тогда станет опять полупроводником *n*-типа по всей оставшейся длине. Эта вторая область кристалла *n*-типа является основой для вырезания эмитерных концов кристаллических элементов транзисторов.

Активирующие таблетки очень малы; размер их примерно равен размеру небольшой булавочной головки. Их теплоемкость незначительна по сравнению с теплоемкостью расплавленного вещества, и введение их в расплав не вызывает заметного изменения температуры или закономерности роста. Часто таблетка сама является кусочком германия, сплавленного с примесью, так как во многих случаях требуемое количество примеси (если взять ее в чистом виде) слишком мало, чтобы с ним было удобно оперировать. Таблетка расплавляется и равномерно распределяется по объему расплава, находящегося в состоянии интенсивного конвективного перемешивания, уже в течение долей секунды.

Во время роста кристалла жидкость осторожно перемешивают. По мере роста кристалла выталкиваемые примеси накапливаются в слоях жидкости, расположенных около границы между твердой и жидкой фазами. Такое накопление примесей приводит в результате к увеличению концентрации в частях кристалла, которые затвердевают позже. Эта концентрация сильно колеблется в результате случайных волн конвекции в жидкости, то совершенно уничтожающих скопление примеси, то снова его создающих в совершенно беспорядочной последовательности. Такие колебания, конечно, нежелательны, поскольку в результате их теряется контроль над местными свойствами кристалла.

Одним из способов обойти эту трудность является очень медленное выращивание кристалла, при котором увеличение концентрации примесей поддерживается пренебрежимо малым вследствие совместного действия диффузии и конвекции. Однако эта скорость роста недопустимо медленна, а расплав постепенно загрязняется примесями из-за абсорбции их из тигля и окружающей атмосферы. Другим, лучшим и фактически единственным применяемым способом является осторожное форсирование конвекции и турбулентности в жидкости, чтобы ускорить выравнивание концентрации в расплаве. Это ускорение движений в жидкости частично обеспечивается геометрией тигля, нагревательной катушки, индуктора и теплопроводов. Кроме того, кристалл вращается вокруг своей вертикальной оси. Это движение создает центробежный рассеивающий эффект в слоях жидкости, соприкасающихся с граничным слоем. Далее, специальное устройство колеблет кристалл в вертикальном направлении. Эти колебания производят всасывание и выталкивание жидкости. Все эти технологические приемы позволяют выращивать кристаллы с однородными электрическими свойствами, достаточно близкими к предварительно рассчитанным.

В современной практике вытягивания кристаллов широко применяется автоматизация. Управляющие устройства поддерживают температуру расплава, регулируют скорость вытягивания и вводят таблетки по заданной программе, включаемой в действие, когда оператор установил контакт между зерном и расплавом и определил, что процесс начался должным образом.

25-3. ВЫРАЩИВАНИЕ С ПЕРЕМЕННОЙ СКОРОСТЬЮ

Накопление примесей перед продвигающейся границей твердой и жидкой фаз и изменение этой повышенной концентрации примесей по мере роста твердой фазы позволяют создать еще один метод выращивания переходов и их последовательного чередования. Этот метод называется «выращиванием с переменной скоростью».

Предположим, что кристалл вытягивается из расплава, содержащего как доноры, так и акцепторы. Предположим далее, что донорная примесь первоначально большой концентрации имеет большую величину K, а акцепторная характеризуется малым K. Если кристалл выращивается методом переменного быстрого и медленного вытягивания, то распределению концентрации доноров и акцепторов вдоль оси кристалла можно придать вид, показанный на рис. 25-6. Концентрация акцепторов в кристалле колеблется сильнее, так как вследствие малости величины K выталкиваемые примеси накапливаются перед растущим кристаллом в периоды быстрого роста в большем количестве и поэтому в твердом веществе они оказываются в большей концентрации. В периоды медленного роста скопившаяся примесь диффундирует в стороны, и распределение этой примеси в расплаве становится более равномерным. Кроме того, средняя концентрация этой примеси по мере роста кристалла возрастает относительно быстрее. С другой стороны, примесь, имеющая большую величину K, проявляет меньшую тенденцию накапливаться перед поверхностью растущего кристалла во время быстрого роста, а поэтому испытывает меньшие колебания концентрации между областями кристалла, выращенными быстро и медленно. Кроме того, средняя концентрация этой примеси медленнее увеличивается от затравочного конца к комлю.

Эти разностные явления проявляются в наибольшей степени, когда конвекция и турбулентное перемешивание доведены до минимума и если кристалл не вращается и не встряхивается. Не рекомендуется, чтобы при выращивании с переменной скоростью процессы накапливания выталкиваемых примесей и диффузии взаимодействовали в результате механической конвекции, тогда как механическое выравнивание концентрации очень желательно при выращивании кристаллов с постоянной скоростью, описанном в предыдущем параграфе.

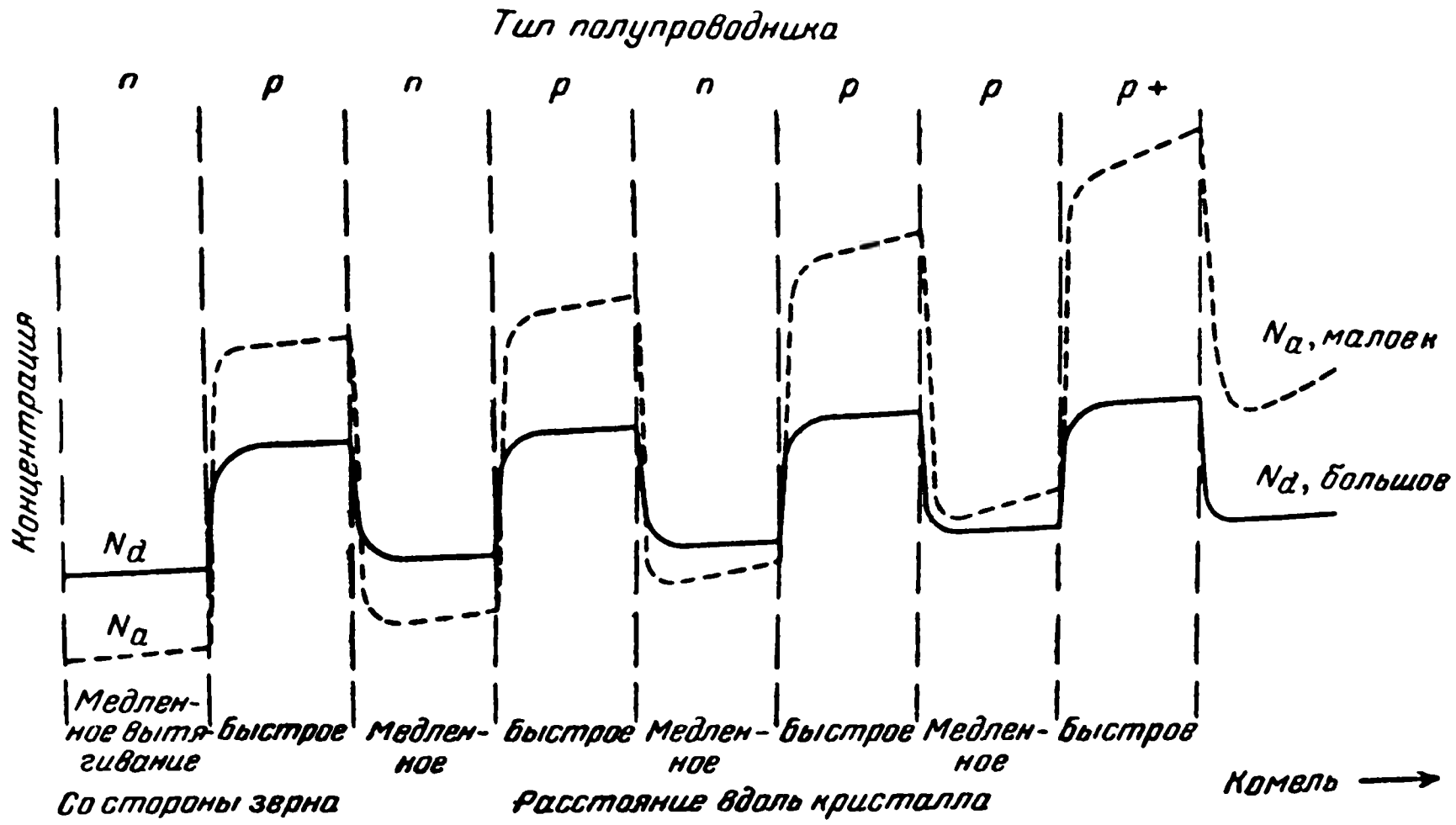


Рис. 25-6 Относительные распределения концентрации доноров и акцепторов в интервалах быстрого и медленного роста для образца, полученного методом вытягивания, видно, как образуются *n*- и *p*-области при изменении быстроты выращивания

Если примеси правильно выбраны, если их первоначальная концентрация в расплаве установлена правильно и если скорости быстрого и медленного выращивания и интервалы между ними правильно подобраны в отношении диффузионных констант примесей в жидкости, то на некоторой части длины кристалла образуется ряд смен типов полупроводника то n-типа, то p-типа.

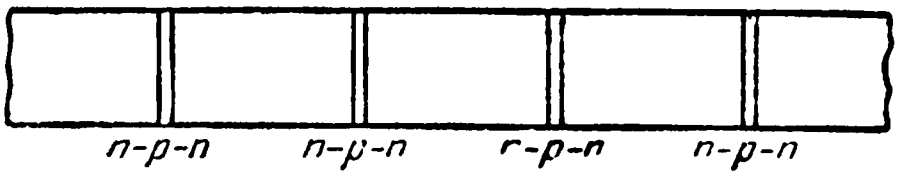


Рис 25-7 В кристалле, полученном методом изменения быстроты выращивания с несколькими структурами переходов типа n-p-n, бо́льшая часть всего объема кристалла пригодна для изготовления транзисторов.

Современная технология выращивания кристаллов с переменной скоростью значительно сложнее, чем изложено в этом кратком описании. Мы только дали представление о природе этого процесса и о его физической и физико-химической основах. Кристаллы, выращенные с переменной скоростью, могут быть вытянуты с несколькими системами n-p-n переходов (рис. 25-7), каждая из которых пригодна для производства многих отдельных элементов с системой переходов n-p-n. С помощью этого метода значительно бо́льшая часть объема всего кристалла становится пригодной для изготовления транзисторов, чем при изготовлении кристаллов методом добавления примесей в таблетках, при котором получается только одна система n-p-n переходов.

25-4. СПЛАВЛЕНИЕ

Процесс сплавления [Л. 8 и 9] для производства p-n переходов и триодов со структурами n-p-n и p-n-p разработан настолько, что с его помощью теперь можно изготовлять ряд типов диодов и триодов.

Термин «сплавление» применяется в его основном значении. Если маленькую таблетку, скажем индия, поместить на верхнюю поверхность германиевой пластинки n-типа и затем нагреть их до температуры, превышающей температуру эвтектики германий — индий,

то в образовавшемся углублении на поверхности образуется маленькая капля расплавленной смеси Ge-In. На фазовой диаграмме Ge-In (рис. 25-8) можно проследить, что́ при этом происходит. Эта диаграмма отличается от обобщенной диаграммы (рис. 25-1) в том отношении, что эвтектическая точка у индиевого конца на рисунке так близка к 100%, что ее нельзя разрешить в масштабе, в котором вычерчена эта

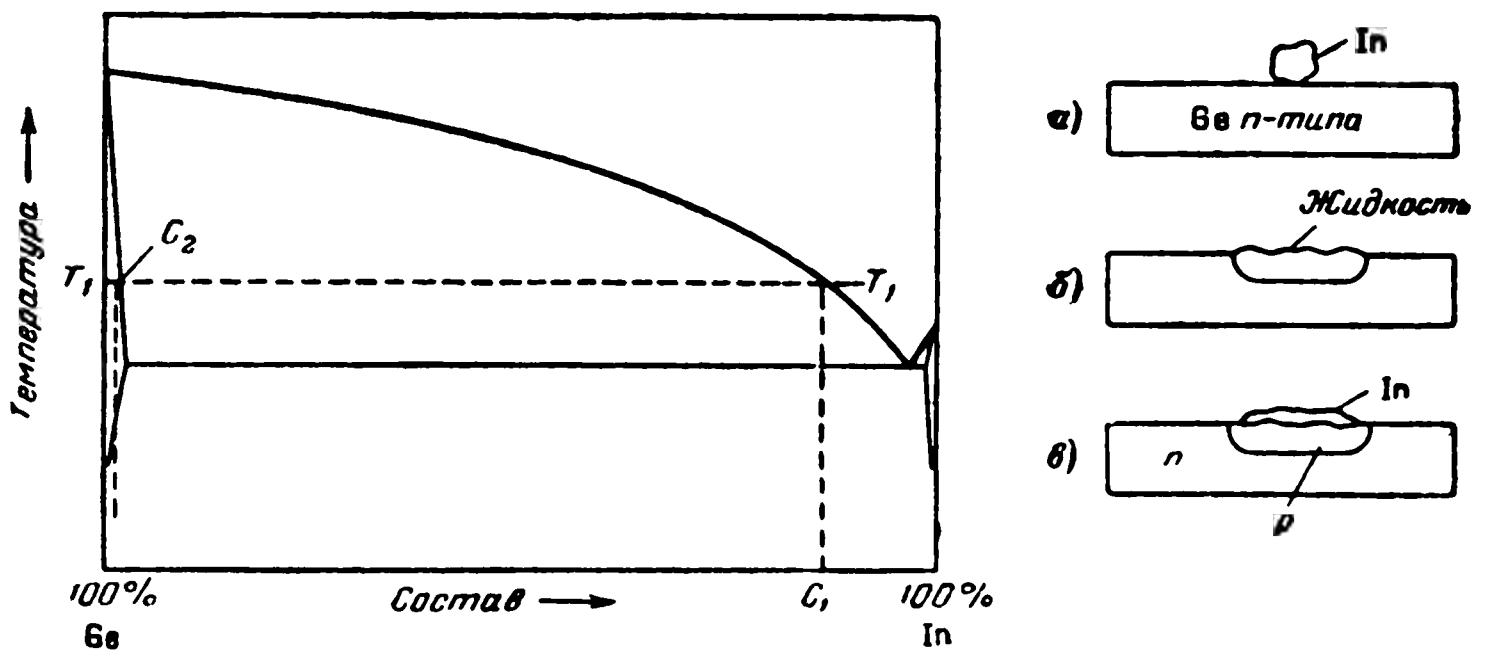


Рис. 25-8 Фазовая диаграмма для системы Ge-In иллюстрирует три стадии получения сплавного перехода

диаграмма [Л. 10]. Рисунок этот несколько искажен, чтобы можно было показать эвтектическую точку. При температуре печи, представленной горизонтальной пунктирной линией T_1, индий плавится и растворяет достаточное количество германия, чтобы получить каплю жидкости состава C_1. Объем жидкости определяется температурой печи T_1 и исходным количеством индия.

При понижении температуры капля начнет затвердевать, и первой отвердевшей порцией будет германий, содержащий индий в концентрации, соответствующей составу C_2 на диаграмме. Этот материал откладывается на твердом германиевом стержне — основе, подобно затравочному зерну при кристаллизации. Содержание индия в составе C_2 соответствует сильной активировке осаждаемого германия акцепторными примесями. По мере хода процесса кристаллизации неотвердевший расплав все более насыщается индием, пока температура не упадет до эвтектической и приктически весь германий не осядет, а оставшийся материал (эвтектика), почти целиком состоящий из индия, не отвердеет в виде метал-

лической пуговицы на внешней поверхности рекристаллизованного участка. Эта пуговица находится в хорошем механическом и невыпрямляющем электрическом контакте с рекристаллизованной областью, и к ней удобно припаивать выводы.

Пока температура описанного печного процесса высока, индий успевает диффундировать в твердое вещество на небольшое расстояние от начальной границы расплавленной фазы. Поэтому полученный *p-n* переход образуется на расстоянии нескольких микрон от границы рекристаллизации.

Повторив эту операцию на противоположной стороне элемента, получим *p-n-p* конструкцию, сходную по геометрии с элементом с *n-p-n* структурой сплавного транзистора, изображенного на рис. 10-14. В производстве обе стороны сплавляются в печи за один цикл после того, как обе пуговицы закрепляются в результате кратковременной термообработки в печке при невысокой температуре. Размеры капель так малы, что они удерживаются на месте во время процесса сплавления силой поверхностного натяжения. Как показано на рис. 10-14, размер коллектора делается несколько бóльшим размера эмиттера. Это обеспечивает больший коэффициент передачи тока в базе с меньшей потерей неосновных носителей вследствие диффузии в стороны из области между переходами.

Если поверхности элементов вырезаны параллельно кристаллографическим плоскостям (*111*), наплавленные участки имеют почти плоские основания по всей площади и получающиеся в результате *p-n* переходы почти параллельны, даже если противоположные поверхности элемента несколько неровны или слегка непараллельны. Физическое объяснение того, что основание оказывается плоским, заключается в том, что плоскости кристалла (*111*) являются самыми плотно упакованными плоскостями в кубическом кристалле и требуется сравнительно большая энергия, чтобы удалить атом из плоскости (*111*), не разрушенной на границе между твердой и жидкой фазами. Однако когда какой-либо атом удален и плоскость нарушена, она очень быстро распадается в расплавленном веществе. Поэтому плавление происходит по всей плоскости одновременно и получающаяся в результате при наивысшей темпера-

туре термообработки в печи наиболее глубоко расположенная поверхность раздела твердой и жидкой фаз должна быть по всему протяжению, по крайней мере в принципе, идеальной атомной кристаллической плоскостью.

25-5. ДИФФУЗИОННЫЕ ПЕРЕХОДЫ

Самым современным и в некотором роде самым многообещающим методом изготовления переходов и структур является диффузионный метод [Л. 11 и 12]. В диффузионном процессе примеси, создающие проводимость одного типа, заставляют диффундировать при высокой температуре в поверхностные слои полупроводника с проводимостью противоположного типа, чтобы создать пленку из обращенного вещества с p-n переходом между этой пленкой и глубже расположенным основным телом. В настоящее время диффузионный процесс применяется для изготовления кремниевых фотоэлектрических солнечных элементов, диодов с большим током инъекции, структур с диффузионной базой и p-n-i-p структур.

При производстве переходов диффузионным методом полупроводниковый элемент (отрезанный, отшлифованный и протравленный до получения желаемого размера и формы) помещается в печь в атмосфере, содержащей газообразную примесь, которую надо ввести диффузией. Эти материалы могут быть запаяны в трубку до обработки в печи, либо диффундирующий пар пропускается над образцом в потоке инертного газа, протекающего через трубку во время термообработки. Время и температура диффузии устанавливаются в зависимости от величины коэффициента диффузии примеси в твердом веществе, от поверхностной концентрации примеси и от желаемой толщины обращенной пленки. Концентрация примеси на любой глубине x под поверхностью выражается через

$$C_x = C_{\text{пов}} \left[1 - \frac{1}{\sqrt{\pi D t}} \int_0^{x/2\sqrt{Dt}} e^{-x^2/4Dt} dx \right]; \qquad (25\text{-}2)$$

$$C_{\text{пов}} \operatorname{erfc} \left(\frac{x}{2\sqrt{Dt}} \right), \qquad (25\text{-}3)$$

где D — коэффициент диффузии примеси в твердом веществе; t — время термообработки.

Температура термообработки входит в это выражение неявно через зависимость коэффициента диффузии от температуры:

$$D = Ae^{-E_{ак}/kT}, \qquad (25\text{-}4)$$

где A и $E_{ак}$ — константы для данной примеси в данном полупроводнике; $E_{ак}$ — энергия активации для процесса диффузии.

На рис. 25-9 приведена зависимость от расстояния концентрации диффундирующей примеси в полупроводнике n-типа. Если, например, примесь является акцептором с поверхностной концентрацией, превышающей кон-

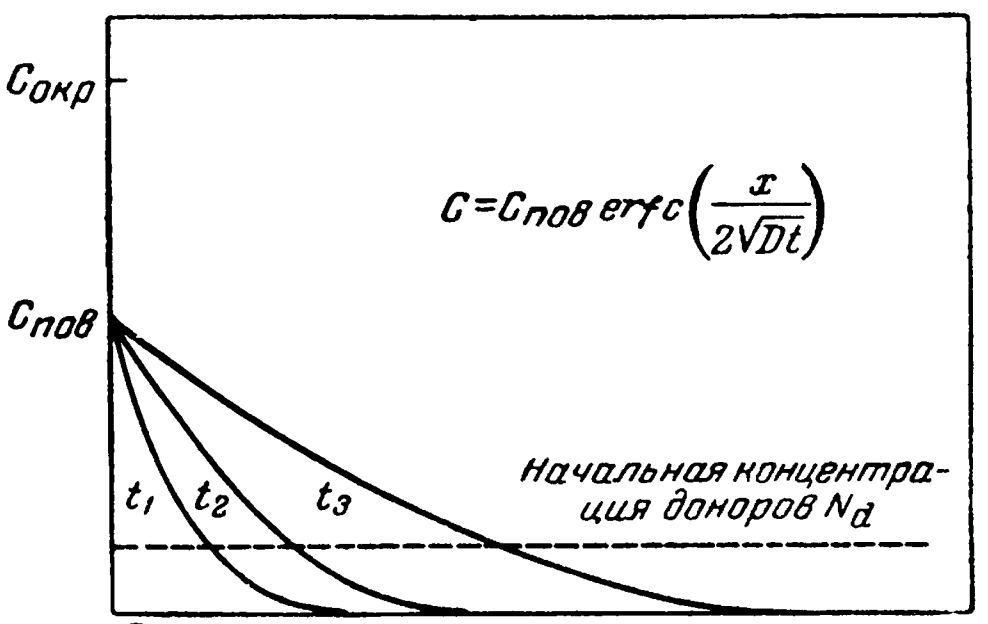


Рис. 25-9 Распределение концентрации примеси, диффундирующей в поверхностный слой полупроводника, зависит от времени и температуры печного цикла

центрацию доноров N_d в полупроводнике, поверхностные слои полупроводника будут содержать больше акцепторов, чем доноров, и будут соответственно p-типа до пересечения кривых концентрации акцепторов и доноров.

Максимальная возможная поверхностная концентрация определяется растворимостью диффундирующего вещества в твердом полупроводнике при температуре термообработки.

В большинстве случаев диффузия происходит при достаточно высокой концентрации примеси в окружающей атмосфере, чтобы $C_{пов}$ в течение всего цикла термообработки было равно этой предельной величине, определяемой растворимостью. В этом случае ее можно зара-

нее определить, что позволяет предвычислить свойства прибора.

Переходы *n-p-n* или *p-n-p* могут быть созданы диффузионным методом путем введения акцепторной примеси с двух сторон образца. Диффузия продолжается до тех пор, пока оба перехода не приблизятся друг к другу на достаточно малое расстояние. Затем края элемента шлифуют, чтобы удалить поверхностный слой, соединяющий эмитерную и коллекторную области в образовавшейся структуре. Небольшая часть эмитерной пленки вытравливается, чтобы открыть исходный слой — базу, к которой затем присоединяется контакт. Иногда для осуществления контакта с базой до проведения диффузионного процесса небольшой участок поверхности исходного элемента покрывается тонким слоем подходящего защитного материала [Л. 14]. Если этот материал удачно выбран и защита правильно выполнена, то доступ примесей к полупроводнику на этом участке будет предотвращен и оголяющаяся после снятия защитного слоя структура готова для присоединения контактов. Тогда геометрия устройства будет такой, как показано на рис. 12-10,*а*.

Производство близко расположенных переходов методом диффузии с противоположных сторон элемента ограничивается требованием параллельности обеих сторон элемента с точностью до долей того конечного расстояния, которое должно разделять переходы. Исходные элементы из германия толщиной 0,125 *мм* можно сделать легко, но гораздо труднее обеспечить параллельность противоположных плоскостей в пределах $\pm 0{,}0125$ *мм*. Эта трудность отсутствует при изготовлении *n-p-n* и *p-n-p* структур методом последовательной или одновременной диффузии доноров и акцепторов с одной стороны элемента.

25-6. ОБЩИЕ ЗАМЕЧАНИЯ

В этой главе мы описали несколько методов технологии полупроводниковых приборов. Мы сделали это достаточно детально, чтобы дать представление о том, что происходит при различных операциях. Кроме того, мы хотели объяснить физическую и химическую природы этих операций. В то же время мы не приводили слишком

много подробностей, которые можно найти в литературе по технологии производства полупроводниковых приборов.

ЗАДАЧИ

1. Коэффициент разделения для галлия равен 0,10. Образец германия с содержанием 0,01% галлия плавится и затем отвердевает. 10% слитка, отвердевшие последними, отрезаются и удаляются, остальное плавится, размешивается до однородности и затем снова медленно замораживается. Затвердевшие последними 10% слитка снова отрезаются и удаляются. Какова средняя концентрация галлия в получившемся образце?

2. В тигле плавится 20 $см^3$ германия первоначально n-типа, имеющего 10^{17} доноров в 1 $см^3$ с коэффициентом разделения 0,05. Из этой расплавленной массы вытягивается кристалл. Допустим, что для обеспечения однородности расплавленная масса перемешивается в течение всего процесса вытягивания. Докажите, что удельное сопротивление отвердевающего материала при комнатной температуре в момент, когда половина расплавленного вещества отвердела, равно 0,17 *ом·см*. В этот момент в расплавленное вещество вводится таблетка галлия, имеющего коэффициент разделения 0,10. Германий, твердеющий в кристалл немедленно после этого, должен иметь удельное сопротивление при комнатной температуре 0,20 *ом·см*. Докажите, что таблетка должна содержать около 0,33 *мг* галлия. При вычислениях считайте, что каждый атом примеси вкладывает в механизм проводимости одну подвижную частицу при комнатной температуре. Атомный вес галлия 70.

3. Образец германия p-типа, имеющий начальное удельное сопротивление при комнатной температуре 1 *ом·см*, подвергается в течение 10 *мин* при температуре 700° С воздействию паров фосфора с концентрацией 10^{18} атомов в 1 $см^3$ в потоке инертного газа при атмосферном давлении. В этих условиях равновесная концентрация фосфора в поверхностном слое германия равна 0,1 концентрации фосфора в газе, находящемся в контакте с поверхностью. Коэффициент диффузии фосфора в германии при этой температуре равен $5 \cdot 10^{-13}$ $см^2/сек$. Вычислите и постройте график удельного сопротивления в функции расстояния от поверхности и докажите, что p-n переход образуется на глубине $5{,}1 \cdot 10^{-5}$ *см*.

4. Базовый слой диффузионного транзистора имеет толщину 0,002 *см*. Концентрация акцепторов в базе спадает от $1{,}1 \cdot 10^{16}$ $см^{-3}$ на эмитерной стороне до 10^{15} $см^{-3}$ на коллекторной стороне. В первом приближении градиент концентрации акцепторов может считаться постоянным, и можно допустить, что акцепторы полностью ионизованы. Вычислите и нанесите на график «вмороженное» поле в базе.

ЛИТЕРАТУРА

1. Pfann W. G., Olsen K. M., Zone-Melting, Bell Laboratories Record, 1955, vol. 33, p. 201; Pfann W. G., Zone-Melting, Wiley, New York, 1958.

2. Pfann W. G., Principles of Zone-Melting, Journal of Metals, 1952, vol. 4, p. 747.

3. Pfann W. G., Redistribution of Solutes by Formation and Solidification of a Molten Zone, Journal of Metals, 1954, vol. 6, p. 294.
4. Czochralski J., Ein neues Verfahren sur Messung der Kristallisationsge—schwindigkeit der Metalle, Zeitschrift für physicalische Chemie, 1918, vol. 92, p. 219.
5. Teal G. K., Little J. B., Growth of Germanium Single Crystals, Physical Review, 1950, vol. 78, p. 647.
6. Teal G. K., Sparks M., Buehler E., Growth of Germanium Single-Crystals Containing p-n Junctions, Physical Review, 1951, vol. 81, p. 637.
7. Hall R. N., p-n Junctions Produced by Growth Rate Variation, Physical Review, 1952, vol. 88, p. 139.
8. Saby J. S., Dunlap W. C., Impurity Diffusion and Space-Charge Layers in 'Fused-Impurity' p-n Junctions, Physical Review, 1953, Vol. 90, p. 630.
9. Ebers J. J., Alloyed Junction Transistor Development, Bell Laboratories Record, 1956, vol. 34, p. 8.
10. Klemm W. et al., Das Verhalten der Elemente der III Gruppe zueinander und zu den Elementen der IV Gruppe, Zeitschrift für anorganische Chemie, 1948, vol. 256, p. 239.
11. Fuller C. S., Struthers J. D., Ditzenberger J. A., Wolfstirn K., Diffusivity and Solid Solubility of Copper in Germanium, Physical Review, 1954, vol. 93, p. 1182.
12. Dunlap W. C., Diffusion of Impurities in Germanium, Physical Review, 1954, vol. 94, p. 1531.
13. Fuller C. S., Ditsenberger J. A., Diffusion of Donor and Acceptor Elements in Silicon, Journal of Applied Physics, 1956, vol. 27, p. 544.
14. Frosch C. J., Derick L., Surface Protection and Selective Masking During Diffusion in Silicon, Electrochemical Society Journal, Dec. 1958.
15. Kromer H., Zur Theorie des Diffusions und des Drifttransistors, Archiv der electrischen Übertragung, 1954, vol. 8, p. 223.
16. Now Extant as Transistor Technology, vol. I, Van Nostrand, Princeton, N. J., 1958.
17. Transistor Technology, vols II and III, Van Nostrand, Princeton, N. J., 1958.